Diabatic Circulations in the Stratosphere.

Steven Pawson

Ph.D.
University of Edinburgh
1988

TABLE OF CONTENTS

Acknowledgements.

I am grateful to the N.E.R.C. for sponsoring this research and the Department of Meteorology at the University of Edinburgh for use of their facilities. Attendance of conferences was made possibility by their generosity and by the Royal Meteorological Society.

In the course of this research I have benefitted from discussions with many people, all of whom have contributed to my understanding of the subject. These include: D.G. Andrews, S.R. Beagley, J.C. Gille, J.D. Haigh, K. Labitzke, D.M. Li, J.C. Marshall, C.J. Marks, M.E. McIntyre, A. O'Neill, V.D. Pope, J.A. Pyle, K. Rose, K.P. Shine and J.F.L. Simmons. I am grateful to Charles Duncan and Keith Weston for their advice and encouragement throughout the course of this work. Most of all, I am indebted to my supervisor, Bob Harwood, for his advice, criticism, encouragement and, most of all his tremendous patience; without him this work would not have been possible.

Finally, I am grateful to my friends and family for their support.

Abstract.

Three dimensional fields of radiative heating and cooling rates have been calculated from satellite observations of temperature, ozone and water vapour distributions in the stratosphere between November 1978 and May 1979. Error analysis shows that these fields are accurate to within <20% over the entire stratosphere. Zonal-mean cross-sections are broadly similar to other recent calculations.

Net radiative heating rates have been obtained from these components; in accord with previous studies it was necessary to impose a constraint on the global heating field. The data were used to show that the mean pressure of each isentropic surface in the stratosphere changes by less than 1% each day, so that the assumption of no net flow of mass through each surface is very accurate; this constraint was applied to the heating fields. Corrected monthly, zonal-mean fields qualitatively agree with the broad range of previous calculations. Three-dimensional heating fields are analysed for the northern hemisphere winter and spring; comparison with the potential vorticity distribution reveals that the strongest net cooling tends to occur at the eastern edge of the displaced polar vortex, consistent with the dynamical situation of warm, tropical air being advected around the vortex by the cyclonic flow.

The net radiative heating rates give the vertical component of the diabatic circulation in isentropic coordinates, since the 'vertical' (cross-isentropic) velocity is directly proportional to the heating rate. The horizontal component of the diabatic circulation is calculated from the vertical mass convergence, for both three-dimensional and zonal-mean cases, using the steady state continuity equation; the satellite data are used to validate the assumption that density changes can be neglected over long times, so that the diabatic circulation is the dominant part of the divergent flow; this is not the case over short periods.

The zonal-mean diabatic circulation agrees well with the most convincing previous calculations; there is ascent in the tropics and descent at the winter pole, with ascent overlying weaker descent at the summer pole, with poleward flow which is strongest near the winter hemisphere stratopause. These circulations have been used with the eddy vertical velocity, determined from

the three-dimensional heating field, to examine various terms in the zonal-mean tracer budget. The cross-isentropic component of transport is found to be dominated by the mean advection, so that the diabatic eddy transport may be neglected. The consequence of this is that the eddy transport may be very accurately approximated by a single, isentropic, diffusion coefficient. Practicalities of zonal-mean modelling in isentropic coordinates are discussed.

The three-dimensional diabatic circulation is compared to the geostrophic flow. Near the stratopause the northward diabatic flow is found to be strongest in regions where the northward geostrophic flow is small; this nongeostrophic flow is then around 20% of the magnitude of the geostrophic flow. There is no such distinct relationship between the two fields in the middle stratosphere (near 30km).

Radiative dissipation coefficients have been calculated for the observed planetary waves in the stratosphere. Use of atmospheric data enables the effects of scale dependence and photochemical effects to be implicitly included in the calculations. In the high stratosphere the dissipation is primarily due to longwave cooling by carbon dioxide; with photochemical acceleration also playing an important role. At lower levels the longwave radiation transfer by ozone and water vapour are important for the dissipation of the waves; evidence suggests that energy exchange in the 9.6μm bands of ozone accelerates the waves, and that in the tropics this is the dominant feature. In middle latitudes the vertical structure of the waves does not significantly affect the dissipation rate, but in the tropics it is important, leading to stronger thermal damping than 'Newtonian' cooling coefficients suggest.

CHAPTER 1

INTRODUCTION: MOTIVATION AND PROBLEMS.

1.1. Introduction.

The middle atmosphere of Earth, the stratosphere and mesosphere, extends from around 10km to near 80km. Zonal-mean cross-sections of the temperature structure (e.g. figure 1.1) show that the temperature minimum of the tropopause is close to 10km in middle latitudes but rather higher, near 15km, in the tropics; the stratosphere is the layer where the temperature increases with altitude, reaching a maximum at the stratopause, which is situated close to 50km in the summer hemisphere and tropics, but slopes upwards and polewards in the winter hemisphere, reaching 55km at the winter pole. In the mesosphere the temperature decreases with height to the minimum at the mesopause (near 80km).

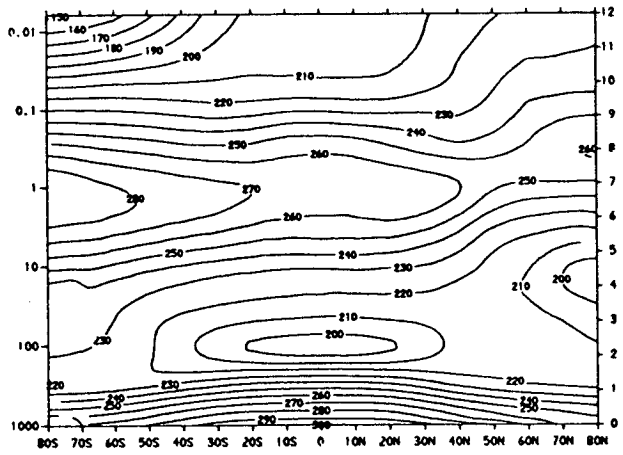


Figure 1.1. The monthly, zonal-mean, climatological temperature structure of the atmosphere in December (from Barnett and Corney, 1985) (contour interval 10K). The right-hand ordinate is the pressure scale-height, $\log(p_o/p)$; when multiplied by 7km this gives the approximate altitude.

It is the statically stable stratosphere which is the topic of this thesis; the reasons for our interest are introduced in this chapter - this necessitates at least some knowledge photochemical processes (which are not treated comprehensively) as well as an understanding of radiative and dynamical processes, which are considered in later chapters. In section 2 of this chapter the distribution of ozone in the atmosphere is discussed; in section 3 possible mechanisms for destroying ozone are introduced and the environmental problems which they can potentially cause are described - this is one of the motivations for studies of the middle atmosphere; finally, section 4 describes the objectives of this thesis.

1.2. The Formation and Distribution of Ozone.

A major motivation for studying the middle atmosphere is the presence of ozone, which shields the surface of Earth from potentially harmful ultraviolet radiation. Ozone acts as a filter of the solar spectrum, absorbing ultraviolet radiation which would otherwise penetrate to the surface of Earth. If this were not the case, severe environmental problems could occur, because the frequencies absorbed by ozone are harmful to living organisms. Direct exposure of human skin to ultraviolet radiation can lead to the development of cancers. Secondary effects are ecologically more complex; for instance, phytoplankton resident near the ocean surface are destroyed by ultraviolet radiation: thus the first link of an ecological chain, involving fish and mammals, is destroyed.

The zonal-mean ozone distribution (figure 1.2) shows that at all latitudes there is a maximum mixing ratio near 30km; also, the maximum occurs in the tropics, where it is usually around 8-12ppmv (depending on season). In order to understand this distribution some simple photochemistry and transport mechanisms must be considered.

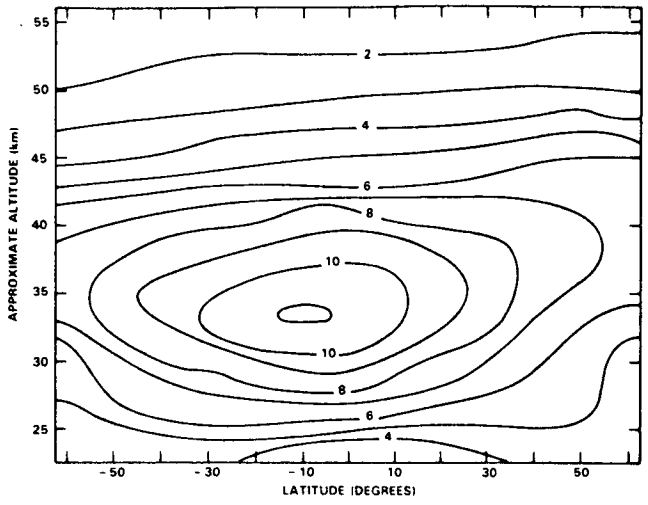


Figure 1.2. Four-year average (1979–1983) monthly, zonal-mean ozone distribution for January (ppmv) observed by the Solar Backscattered Ultraviolet Radiometer (contour interval 1ppmv). From WMO (1986).

Ozone (O_3) is a photochemically active gas, produced by the interaction of a diatomic oxygen molecule (O_2) with an oxygen atom (O) and destroyed by photodissociation (energy is provided by a photon). The simplest model of ozone photochemistry is the Chapman (1930) scheme, which may be expressed symbolically as:

$$O_2 + h\nu_1 \rightarrow 2O \quad \text{(slow)} \tag{1.1}$$

$$O_2 + O + M \rightarrow O_3 + M \quad \text{(fast)} \tag{1.2}$$

$$O_3 + h\nu_2 \rightarrow O + O_2 \quad \text{(fast)} \tag{1.3}$$

$$O + O_3 \rightarrow 2O_2 \quad \text{(slow)} \tag{1.4}$$

Here h is Plancks constant, ν_1 and ν_2 are two different frequencies of solar (ultraviolet) radiation, so that $h\nu_1$ and $h\nu_2$ represent different quanta of energy, and M represents a 'third body', a molecule or atom which will undergo a three-body collision without reacting with the oxygen species.

Reactions (1.1) and (1.4) are 'slow' whilst the others are 'fast'; thus, to a first approximation the relative abundances of O, O_2 and O_3 are determined by the reaction rates of reactions (1.2) and (1.3). Because density decreases with height the probability of three-body collisions diminishes with increasing altitude; the number of photons in the solar beam with frequency ν_2 decreases by absorption as the beam passes downwards through the atmosphere; allowing also for the increase in oxygen atoms with increasing height (reaction (1.1)) this suggests that there is likely to be a level of maximum ozone mixing ratio, which will depend upon latitude and season. Such a distribution indeed calculated by simple photochemical equilibrium models (where the equilibrium ozone distribution for a given solar illumination is calculated). However, photochemical equilibrium ozone distributions differ from observations; in the global-mean, too much ozone is predicted, so that other destruction mechanisms must operate (as discussed below); also, the tropical ozone maximum is situated too low and too little ozone is calculated in middle and high latitudes. To explain these discrepancies other mechanisms must be invoked: the most obvious of these is that ozone is transported away from its source region in the tropical stratosphere to higher levels and middle latitudes.

Consider a zonal-mean viewpoint, so that there are two possible transport mechanisms, as will be discussed more fully in chapters 2 and 6; these are planetary waves and the mean-meridional-circulation. Both are consistent with the differences between the observed ozone distribution and the photochemical equilibrium distributions, as is now discussed.

Observations of nonlinear planetary waves in the winter stratosphere (e.g. McIntyre and Palmer, 1983) show that air of tropical origin is transported into middle and high latitudes in the winter stratosphere. Further, observations of the ozone distribution near 10mbar (Leovy *et al*, 1985) reveal that these eddies also transport ozone. Thus, at least qualitatively these waves can explain the meridional distribution of ozone: they transport ozone almost horizontally from its equatorial source region to middle and high latitudes.

Imbalance between the solar heating and terrestrial cooling (to be discussed more fully in chapters 4 and 5) leads to regions of net heating and cooling in the middle atmosphere; generally there is net heating at all levels in

the tropical stratosphere and cooling in middle and high latitudes of the winter hemisphere. If, after Dunkerton (1978), we associate ascending motion with regions of net radiative heating, there must be upward motion in the tropics, which would transport ozone vertically; this can qualitatively explain why the tropical ozone maximum is observed to be higher than simple chemical models predict.

An important factor in our understanding of the middle atmosphere is an appreciation of the inter-relationship of dynamics, radiation and photochemistry. Ozone is transported away from its source region to higher levels by the diabatic vertical velocity; this itself depends on the strength of solar heating in the tropics, which is dependent upon the absorption of photons by ozone. Thus, an increase in the ozone distribution can lead to increased heating (neglecting the increased cooling in the ozone bands) which enhances the vertical transport, thus transporting ozone more rapidly from its source region. Similarly, the meridional transport of ozone by planetary waves changes the radiative properties of middle latitudes, which affects the temperature structure and the propagation of the planetary waves themselves. A complete understanding of the middle atmosphere thus requires not only a knowledge of the individual processes, but also an understanding of their interactions. This thesis is an attempt to understand some aspects of the radiative-dynamical coupling in the stratosphere.

As yet, the major reasons for study of the ozone layer have not been put forward; this is dealt with now.

1.3. The Threat to the Ozone Layer.

In the Chapman scheme for the ozone budget, reaction (1.4) is the only mechanism by which ozone can be destroyed. In practice this destruction is too slow to destroy ozone at a rate consistent with observations of the ozone distribution. A variety of other chemical species in the stratosphere are responsible for the destruction of ozone; some of the most significant are the catalytic cycles involving hydroxyl (OH), Chlorine (Cl) and the oxides of nitrogen (NO_x). The significance of these being catalytic reactions is that after they are complete the cycle can recommence, so that ozone can be continually destroyed. Thus, if these chemical reactions can occur on a timescale much more rapid than ozone can be replenished by dynamical or

other chemical mechanisms, it is potentially possible to remove a large part of the ozone from a region of the atmosphere. Photochemical details of such reactions are discussed by Brasseur and Solomon (1986).

Because there are no chemical cycles which provide a long-term reservoir for chlorine in the middle atmosphere and since it has a very long residence time (determined by how rapidly it may be transported into the troposphere or escape to space), any source of the gas causes an accumulation of it in the middle atmosphere.

There are anthropogenic sources of chlorine. Chlorofluorocarbons (CFCs) are used as propellants in aerosol cans, coolants in refrigeration systems and are a bi-product in the production of some expanded polystyrenes; all result in escape of CFC's into the atmosphere. Once these gases are in the troposphere they can be transported into the stratosphere where the chlorine is released by photolysis. In the early 1970's anthropogenic activity led to a release of CFC's sufficient to quadruple the amount of chlorine in the atmosphere by the middle of next century. Because of its catalytic role in ozone destruction, its increased abundance in the middle atmosphere could have a potentially disastrous effect on the ozone layer; this problem has motivated many studies of the middle atmosphere which attempt to predict the effect of chlorine on the ozone layer.

A further anthropogenically-induced perturbation to the middle atmosphere is the large input of CO_2 arising from the combustion of fossil fuels. Estimates from the early 1970's predicted that the CO_2 content of the atmosphere would double by 2040, although the fossil fuel consumption rate has decreased since this estimate was made. Because the ozone distribution is temperature dependent, particularly in the high stratosphere, the change in temperature caused by the changes in radiative balance brought about by the increased CO_2 content are likely to lead to changes in the O_3 distribution. This too has been the subject of modelling studies.

Thus the major motivation for study of the middle atmosphere is that anthropogenic activity can lead to changes in the ozone distribution: specifically, ozone can be destroyed. The extent of ozone destruction poses a major environmental problem; reliable assessments of its possible extent are required to assist with worldwide environmental policy making without which

the problem could evolve unchecked.

The consequences of CFC emission are already evident. Farman *et al.* (1985) noted that in successive Antarctic springs since 1979 the total ozone column amount decreased. Further studies, including special observation programs of the Antarctic atmosphere in 1986 and (as yet unreported) 1987, have shown that almost all of the ozone disappears from the low stratosphere in the early spring and that the cause of this ozone destruction appears to be caused by chemical destruction involving Chlorine.

1.4. The Scope of this Thesis.

This thesis is primarily concerned with the diagnosis of diabatic circulations in the stratosphere between November 1978 and May 1979, using observations of the temperature, ozone and water vapour distributions by the Limb Infrared Monitor of the Stratosphere (LIMS) experiment. As such, it represents a limited study of some radiative and dynamical aspects in the stratosphere; interpretation of these calculations should lead to an increased understanding of the stratosphere which can be used to isolate some of the important processes from the unimportant ones. This knowledge can then be used to help simplify the formulation of numerical models and can also assist in the interpretation of other data sets which may not be as comprehensive as LIMS.

The relevant theory concerning diabatic circulations and their importance in the middle atmosphere is outlined in chapter 2. The first step in obtaining the diabatic circulation is the determination of the net radiative heating rate; the radiation transfer model used to this end is discussed in chapter 4 following a description of the data used (chapter 3). Also discussed in chapter 4 are some attempts to estimate the accuracy of the radiative heating and cooling rates given the uncertainty in the data; this determines the confidence we can place in the results of later chapters.

Net radiative heating rates, obtained from the sum of solar and terrestrial components, are introduced in chapter 5. In accord with previous studies, simply adding the heating and cooling rates does not give a consistent energy budget; the net heating rate must be adjusted for consistency. Use of isentropic coordinates enables the imposition of a simple, but entirely

consistent constraint on the the net heating field. The adjusted fields are presented for each month: zonal-mean distributions describe the latitude-height variations, whilst polar stereographic plots enable the relationship between the heating distribution and the vortex structure to be determined.

In chapter 6 some problems relevant to zonal-mean modelling are discussed. The zonal-mean diabatic circulation for each month of LIMS data is presented; these can be compared to previous studies and also serve as a reference for any future modelling studies. An assessment is made of various terms in the zonal-mean tracer budget equation; the cross isentropic transport of trace gases by mean and eddy motions is considered. The results of this work enable justification of various approximations to the zonal-mean tracer budget in isentropic coordinates, which considerably simplify the transport equation. Implications are discussed.

Three-dimensional divergent circulations are discussed in chapter 7. In the monthly-mean the diabatic circulation is found to be the dominant part of the divergent flow. The structure of the diabatic flow in comparison to the geostrophic flow is discussed and implications deduced. No attempt is made to perform a quantitative analysis of the influence of the diabatic flow on the evolution of the rotational flow; the analysis is error-prone.

In chapter 8 the observed eddy distributions of radiative heating and cooling have been examined and used to calculate radiative dissipation coefficients for these waves in the stratosphere. These results are more relevant than the 'Newtonian cooling' often used in mechanistic studies of wave propagation in the middle atmosphere, and include a degree of scale dependence, which is most important in the tropical stratosphere.

Finally, chapter 9 contains a discussion of the major results of the thesis and possible future, related work.

CHAPTER 2

THE DYNAMICAL FORMULATION AND THE DIABATIC CIRCULATION.

2.1. The Primitive Equations.

2.1.1. The Basic Assumptions.

The assumptions involved in deriving the primitive equations from the complete form of the differential relationships governing the evolution of fluid flow on a rotating planet are discussed by Lorenz (1967) and Dutton (1976); they are not examined in detail here, but summarised. We assume that we are considering a rapidly rotating, shallow, compressible fluid which is in hydrostatic balance; thus the vertical distribution of pressure, p, is related to the geopotential height, z, by:

$$\partial p/\partial z = -\rho g, \tag{2.1}$$

where $g=9.81 ms^{-2}$ is the acceleration due to the gravitational field of Earth and ρ is the density of the atmosphere. It is also assumed that the perfect gas law,

$$p = R\rho T, \tag{2.2}$$

is obeyed; here T represents the temperature and $R=287.05 Jkg^{-1}K^{-1}$ is the gas constant for air. (Note that R is constant only below z=80km where the composition of the atmosphere is uniform; above this altitude the ionisation of various gases leads to a significant change in composition.)

Other assumptions are:

- the radial distance from the centre of Earth is replaced by the radius of Earth, a, except where it is differentiated.
- the vertical component of Earth's rotation dominates the horizontal component.
- gravitational acceleration is assumed constant (and the geopotential height is used rather than the actual altitude).

Lorenz (1967) considers these approximations more fully, outlining how the primitive equations retain the invariant properties of the motion which are possessed by the unapproximated equations.

It is convenient to introduce the concept of potential temperature, θ; this is the temperature an air parcel would have if brought adiabatically to the standard pressure, p_o=1000mbar (1mbar=100Pa). Thus:

$$\theta = T(p_o/p)^{\kappa}, \tag{2.3}$$

where $\kappa=R/c_p=2/7$, $c_p=1005 Jkg^{-1}K^{-1}$ being the specific heat capacity at constant pressure. Note that in terms of potential temperature the entropy is defined as $c_p \log\theta$.

2.1.2. The Navier-Stokes Equations in Isentropic Coordinates.

For the analysis of atmospheric motions it is often convenient to use a vertical coordinate other than geometric height. Assuming hydrostatic balance, transformations to alternative vertical coordinates are straightforward (e.g. Dutton, 1976). The most commonly used vertical coordinates are functions of pressure or entropy; the analysis for this thesis has mostly been performed in isentropic coordinates (for reasons which will become apparent). The equations of motion with $\eta=\log\theta$ as vertical coordinate are now introduced. On occasions logarithmic pressure coordinates will be used (especially in chapters 4 and 8): the relevant equations will be introduced as necessary.

a. The Thermodynamic Equation.

The advantage of using isentropic coordinates (and also the reason for using η rather than θ) as vertical coordinate, is evident when we consider the thermodynamic equation:

$$w = D\eta/Dt = J/(c_pT). \tag{2.4}$$

This equation shows that η changes only in the presence of diabatic heating: in isentropic flow w vanishes, so that the fluid flow is quasi-horizontal. Because of the Lagrangian nature of the vertical coordinate but the retention

of Eulerian coordinates for the horizontal position, isentropic coordinates are known as quasi-Lagrangian. The quasi-Lagrangian nature of the coordinates will be exploited in this thesis; equation (2.4) shows that from a knowledge of the heating rate, J, we can directly deduce the vertical 'velocity', w: this means that the vertical motion in isentropic coordinates is thermally direct. Such a simple relationship is not possible in other coordinate systems, where it is generally necessary to use an ω-equation method to determine the vertical velocity from the distribution of horizontal divergence. Note that this vertical velocity, w, does not include the vertical movement of air parcels associated with undulations in the height of isentropic surfaces; however, this is essentially reversible (the global mean height of isentropic levels is approximately constant). Consequently, any systematic vertical transport of air parcels can only be achieved by the existence of a net diabatic heating rate.

b. The Equation of Continuity of Matter.

The density in isentropic coordinates is related to the density in (λ,ϕ,z)-space as:

$$\sigma_\eta = \rho(\partial z/\partial \eta) = -g^{-1}(\partial p/\partial \eta) \tag{2.5}$$

where the hydrostatic relation, (2.1), has been used in the second part of this relationship.

The equation expressing continuity of matter in isentropic coordinates is:

$$\frac{\partial}{\partial t}(\sigma_\eta) + \nabla_\eta \cdot \mathbf{U} + \frac{\partial W}{\partial \eta} = 0, \tag{2.6}$$

where $\mathbf{U}=\sigma_\eta \mathbf{u}$ and $W=\sigma_\eta w$ are the horizontal and vertical fluxes of matter. The del-operator, ∇_η, is assumed to be at constant entropy.

c. The Hydrostatic Relation.

The Montgomery potential, M, is defined as the sum of the geopotential height and the specific enthalpy:

$$M = c_pT + gz. \tag{2.7}$$

The hydrostatic relation in isentropic coordinates equates the vertical gradient of Montgomery Potential with the temperature distribution:

$$\partial M/\partial \eta = c_pT = c_p\theta\left(p/p_o\right)^{\kappa} \tag{2.8}$$

(the definition (2.3) has been used here).

d. The Horizontal Momentum Equation.

The horizontal momentum equation is:

$$\frac{\partial \mathbf{u}}{\partial t} + \mathbf{u}\cdot\nabla_\eta\mathbf{u} + w\frac{\partial \mathbf{u}}{\partial \eta} + f\mathbf{k}\times\mathbf{u} + \nabla_\eta M = \mathbf{X} \tag{2.9}$$

where $\mathbf{u}$ is the horizontal velocity, $f=2\Omega\sin\phi$ is the Coriolis parameter ($\Omega=7.29\times10^{-5}s^{-1}$ being the rotation rate of Earth) and $\mathbf{X}=(X,Y)$ represents any external forcing per unit mass (and is usually generalised to include molecular viscosity and small-scale phenomena such as dissipating gravity waves). An alternative manner of expressing equation (2.9) is to use the identity $\mathbf{u}\cdot\nabla_\eta\mathbf{u} = \zeta\mathbf{k}\times\mathbf{u} + \nabla_\eta\left(\mathbf{u}\cdot\mathbf{u}/2\right)$, where $\zeta=\mathbf{k}\cdot\nabla_\eta\times\mathbf{u}$ is the relative vorticity, so that:

$$\frac{\partial \mathbf{u}}{\partial t} + w\frac{\partial \mathbf{u}}{\partial \eta} + (\zeta+f)\mathbf{k}\times\mathbf{u} + \nabla_\eta\left(M+\mathbf{u}\cdot\mathbf{u}/2\right) = \mathbf{X}. \tag{2.10}$$

The relevance of this will become apparent later.

2.1.3. Potential Vorticity.

Analysis of atmospheric motion is often simplified by the use of (quasi-) conserved quantities, such as the entropy discussed above. Ertel (1942) showed that the quantity 'potential vorticity' is also conserved when the motion is adiabatic and conservative. Here we consider the isentropic formulation of this relationship, including the nonconservative terms; it can be shown that Ertel's derivation (using the primitive equations) reduces to this form under the coordinate transformation.

The derivation outlines here closely follows Starr (1948). The vertical component of the CURL of the horizontal momentum equation gives the vorticity equation; this is most straightforward if the version (2.10) is used. It is:

$$\frac{\partial \zeta}{\partial t} + w\frac{\partial \zeta}{\partial \eta} + \mathbf{k}\cdot\nabla_\eta w \times \frac{\partial \mathbf{u}}{\partial \eta} + \mathbf{u}\cdot\nabla_\eta(\zeta+f) + (\zeta+f)\nabla_\eta\cdot\mathbf{u} = \mathbf{k}\cdot\nabla_\eta\times\mathbf{X}$$

(we have made use of the vanishing of the curl of a gradient). Noting that the planetary vorticity, f, does not vary with altitude or time, we may re-express the vorticity equation as:

$$D_\eta(\zeta+f) = \mathbf{k}\cdot\nabla_\eta\times\mathbf{X} - w\frac{\partial}{\partial \eta}(\zeta+f) - (\zeta+f)\nabla_\eta\cdot\mathbf{u} - \mathbf{k}\cdot\nabla_\eta\times\frac{\partial \mathbf{u}}{\partial \eta}. \qquad (2.11)$$

Here the absolute derivative following the horizontal velocity on an isentropic surface has been written $D_\eta \equiv (\partial/\partial t + \mathbf{u}\cdot\nabla_\eta)$.

The potential vorticity equation is obtained by subtracting $\sigma_\eta^{-2}(\zeta+f)$ times the continuity equation (2.6) from σ_η^{-1} times equation (2.11), leaving:

$$D_\eta Q = \sigma_\eta^{-1}\mathbf{k}\cdot\nabla_\eta\times\mathbf{X} - w\frac{\partial Q}{\partial \eta} + Q\frac{\partial w}{\partial \eta} - \sigma_\eta^{-1}\mathbf{k}\cdot\nabla_\eta\times\frac{\partial \mathbf{u}}{\partial \eta}, \qquad (2.12)$$

where

$$Q = \sigma_\eta^{-1}(\zeta+f) \qquad (2.13)$$

is the potential vorticity. In the absence of external forcing and diabatic heating the right hand side of equation (2.12) vanishes, showing that the potential vorticity is conserved following the quasi-horizontal flow on the isentropic surface.

An interesting variant of the potential vorticity equation was discovered by Haynes and McIntyre (1987); this is a flux-type theorem which includes the

non-conservative and diabatic terms in the definition of a horizontal vector. This vector, which we shall call the generalised potential vorticity flux, is defined in isentropic coordinates as:

$$\mathbf{\Pi} = \sigma_\eta \mathbf{u} Q + \mathbf{k}\times w\frac{\partial \mathbf{u}}{\partial \eta} - \mathbf{k}\times \mathbf{X}. \tag{2.14}$$

Note that in in component form it may be written

$$\begin{aligned}\mathbf{\Pi} &= (\Pi_\lambda, \Pi_\phi) \\ &= \left(\sigma_\eta u Q - w(\partial v/\partial \eta) + Y,\ \sigma_\eta v Q + w(\partial u/\partial \eta) - X\right)\end{aligned} \tag{2.15}$$

Using equations (2.13) and (2.14) the horizontal momentum equation (2.10) can be written

$$\left(\partial \mathbf{u}/\partial t\right) + \mathbf{k}\times\mathbf{\Pi} + \nabla_\eta\left(M + \mathbf{u}\cdot\mathbf{u}/2\right) = 0. \tag{2.16}$$

This in itself is interesting, since it shows that at steady state (or in the time-mean: D.G. Andrews, personal communication, 1987) the flux of potential vector $\mathbf{\Pi}$ is directed parallel to contours of constant $M+\mathbf{u}\cdot\mathbf{u}/2$, and has the same magnitude as this term.

The flux form of the potential vorticity equation is most easily obtained by considering the divergence of $\mathbf{k}\times$ equation (2.16), and is:

$$\frac{\partial}{\partial t}\left(\sigma_\eta Q\right) + \nabla_\eta\cdot\mathbf{\Pi} = 0. \tag{2.17}$$

The physical significance of this relationship is that it includes the diabatic and nonconservative terms in the definition of potential vorticity, showing that their presence is equivalent to the existence of a horizontal flux of potential vorticity at the isentropic level. It seems simple to consider it in terms of the flux cutting the surface of constant entropy, in analogy with electrodynamical fluxes. In particular, Haynes and McIntyre note that this relationship shows that this generalised potential vorticity flux cannot be transported across isentropic surfaces like a conventional tracer, but merely redistributed at isentropic

layers. Relationship (2.17) shows that only when Π is divergent can the local potential vorticity density change: a rotational flux does not change the flow.

2.1.4. Geostrophic and Thermal Wind Balance.

In isentropic coordinates the geostrophic balance is between a horizontal wind field and the gradient of the Montgomery Potential:

$$u_G = -f^{-1}k\times\nabla_\eta M. \tag{2.18}$$

In common with the pressure- or height-formulation of geostrophic balance this relationship is valid when the Rossby number, which can be expressed as the ratio between the relative and planetary vorticities, is small; then the absolute acceleration may be neglected in the horizontal momentum equation (2.9) (and frictional and diabatic terms are also omitted). One interesting point is that in isentropic coordinates the neglect of the vertical (cross-isentropic) advection of horizontal momentum involved in the geostrophic approximation is exact if there is no diabatic heating, whereas with other vertical coordinates air parcels can have a vertical velocity, even in adiabatic flow, because of the undulation of isentropic surfaces.

Thermal wind balance expresses the vertical gradient of geostrophic wind in terms of the temperature distribution. It is obtained by using the hydrostatic relation (2.8) in the vertical derivative of the geostrophic relation (2.18). It is:

$$f\left(\partial u_G/\partial\eta\right) = -c_p k\times\nabla_\eta T. \tag{2.19}$$

Although strictly applicable to the geostrophic wind, thermal wind balance is often regarded as being well satisfied by the real wind.

2.2. The Zonally-Averaged Equations of Motion.

2.2.1. Definition of the Zonal Average.

The zonal-mean of any quantity s is defined as its average around a latitude circle:

$$[s] = (2\pi)^{-1}\int_0^{2\pi} s\, d\lambda. \tag{2.20}$$

Longitudinal deviations from the zonal-mean are denoted by the prime and defined by:

$$s'(\lambda,\phi,\eta,t) = s(\lambda,\phi,\eta,t) - [s(\phi,\eta,t)]. \qquad (2.21)$$

The eddy term represents the zonally-asymmetric component of the distribution of s.

As pressure is determined by the weight of fluid above any particular height, there are no variations in the mass of fluid contained in each unit volume around the latitude circle in pressure coordinates; thus [s] gives a physically realistic impression of the 'amount' of substance s present in such coordinates. However, the density in isentropic (and height) coordinates can vary with longitude; thus, for some quantities (such as trace gas mixing ratios) it is more meaningful to consider density-weighted zonal-average defined by:

$$\langle s \rangle = [\sigma_\eta s]/[\sigma_\eta], \qquad (2.22)$$

which allows for such longitudinal variations. Gallimore and Johnson (1981) noted that this mean is related to the conventional zonal-mean, equation (2.20), by:

$$\langle s \rangle = [s] + [\sigma_\eta' s']/[\sigma_\eta]. \qquad (2.23)$$

Thus, the mass-weighted zonal-mean contains a component which would conventionally be regarded as an eddy-correlation term.

Deviations from the mass-weighted zonal-mean are defined by:

$$s^* = s - \langle s \rangle \qquad (2.24)$$
$$= s' - [\sigma_\eta' s']/[\sigma_\eta];$$

the relationship to the conventional Eulerian zonal-mean was obtained using the definitions (2.23) and (2.21) in equation (2.24).

Tung (1986) noted that although [s*] does not vanish, the identity

$$[\sigma_\eta s^*] \equiv 0 \tag{2.25}$$

holds.

For future reference we note the following expansions:

$$[\sigma_\eta rs] = [\sigma_\eta]<r><s> + [\sigma_\eta r^* s^*], \tag{2.26}$$

$$[r^* s'] = [r's'], \tag{2.27}$$

where r is any variable, like s.

2.2.2. The Zonally-Averaged Equations.

a. The Continuity Equation.

The zonal-mean of the continuity equation (2.6) is obtained by using the definition (2.22) of the mass-weighted zonal-mean to express $[\sigma_\eta v]=[\sigma_\eta]<v>$ (and similarly for w), to obtain:

$$\frac{\partial}{\partial t}[\sigma_\eta] + \frac{1}{a\cos\phi}\frac{\partial}{\partial \phi}\left([\sigma_\eta]<v>\cos\phi\right) + \frac{\partial}{\partial \eta}\left([\sigma_\eta]<w>\right) = 0. \tag{2.28}$$

Note here that an advantage of using the mass-weighted zonal-mean is immediately apparent, since if ([v],[w]) were used in place of (<v>,<w>) there would also need to be eddy-correlation terms in the analogous equation to (2.28). In the current formulation this eddy contribution is implicitly included in the definition of the meridional flow.

b. The Zonally-Averaged Vertical Velocity.

The mass-weighted zonal-mean vertical velocity is obtained from the zonal-mean of the product of density and the thermodynamic equation (2.4), which gives:

$$[\sigma_\eta]<w> = [\sigma_\eta J/(c_p T)]. \tag{2.29}$$

Using our three-dimensional data we will be able to calculate the right-hand-side of this relationship explicitly. Tung (1982) noted the approximate form of

equation (2.29):

$$[\sigma_\eta]\langle w\rangle \simeq [J]/(c_p\Gamma_e)$$

where $\Gamma_e = (T_e)^{-1}(\partial\eta_e/\partial z)$ is the static stability of the zonal-mean, radiative equilibrium atmosphere (the subscript 'e' has been used to denote this state). This approximation neglects longitudinal variations of density in the atmosphere, as well as assuming that the zonal-mean atmosphere does not depart significantly from radiative equilibrium.

c. The Zonally-Averaged Zonal Momentum Equation.

Starting with the zonal component of equation (2.16), which relates the change in the zonal velocity to the northward generalised potential vorticity flux and the zonal gradient of $M+\mathbf{u}\cdot\mathbf{u}/2$, it is evident that:

$$\partial[u]/\partial t + [\Pi_\phi] = 0. \qquad (2.30)$$

Hence, zonal acceleration, $\partial[u]/\partial t$, is associated with a non-vanishing zonal-mean northward generalised potential vorticity flux (Haynes and McIntyre, 1987). This extremely simple concept of expressing the acceleration of the zonal-mean flow leads to what is perhaps the most fundamental statement of a non-acceleration theorem (Eliassen and Palm, 1961; Charney and Drazin, 1961; Andrews and McIntyre, 1976, 1978): if the mean northward generalised potential vorticity flux vanishes there can be no acceleration of the zonal-mean zonal flow; conversely, a steady zonal-mean zonal velocity implies that $[\Pi_\phi]\equiv 0$. We discuss this further.

This concept contains, as approximate cases, many of the results of Andrews (1983), and their generalisation to diabatic flows by Tung (1986), concerning wave - mean-flow interactions in isentropic coordinates. To see this, write $[\Pi_\phi]$ in terms of its mean and eddy components, which are, respectively:

$$[\Pi_\phi]_M = [\sigma_\eta]\langle v\rangle[Q] - [X] + \langle w\rangle\left(\partial[u]/\partial\eta\right)$$
$$[\Pi_\phi]_E = [(\sigma_\eta v)'Q'] + [w^*]\left(\partial[u]/\partial\eta\right) + [w'\left(\partial u'/\partial\eta\right)]. \qquad (2.31)$$

(equation (2.27) has been used to simplify the latter part of this relation). The eddy correlation terms of $[\Pi_\phi]_E$ can be expressed as the divergence of a vector in the meridional plane; this vector is the Eliassen-Palm flux, whose direction and magnitude indicate the propagation vector for the wave and whose divergence represents a body force on the zonal-mean flow, leading to acceleration (Andrews and McIntyre, 1976, 1978; Andrews, 1983; Tung, 1986). Detailed analysis of the wave propagation is beyond the scope of this thesis: the potential vorticity viewpoint is adequate for future discussion.

Using the definition (2.31) the steady state zonal-mean zonal momentum equation expresses a balance between the mean and eddy northward parts of the generalised potential vorticity flux:

$$[\Pi_\phi]_M = [\Pi_\phi]_E. \tag{2.32}$$

For adiabatic, conservative motion this reduces to:

$$[\sigma_\eta]\langle v\rangle[Q] = -[(\sigma_\eta v)'Q'], \tag{2.33}$$

which should be compared with the quasi-geostrophic Taylor relation in log-pressure coordinates:

$$f[v] = -\rho[v'q']$$

where q is the quasi-geostrophic potential vorticity. Thus, equation (2.33) is the isentropic-coordinate equivalent of the quasi-geostrophic Taylor relationship (note that $[\sigma_\eta][Q]=[\zeta+f]$ and that the relative vorticity is negligible in comparison to the planetary vorticity under the assumption of small Rossby number); further, equation (2.32) extends the Taylor relationship to a general diabatic, non-conservative flow.

d. The Zonally-Averaged Meridional Momentum Equation.

Zonally-averaging the meridional component of the horizontal momentum equation (2.16) gives:

$$\partial[v]/\partial t + [\Pi_\lambda] + a^{-1}[\partial M/\partial\phi] = 0$$

Tung (1986) suggested that the dominant balance in this relationship is:

$$([\zeta] + f)[u] + a^{-1}(\partial[M]/\partial\phi) = 0;$$

that is, the product of absolute zonal-mean vorticity and zonal-mean zonal velocity balances the northward gradient of Montgomery potential.

2.3. The Zonally-Averaged Diabatic Circulation.

2.3.1. The Concept of Radiative Equilibrium.

The temperature and gaseous structure of the atmosphere determine the absorption and emission of electromagnetic radiation. This will be discussed in more detail in chapter 4; at this stage it is important to note that the radiation transfer is not wholly determined by the local state of the atmosphere but depends upon the vertical structure since solar radiation is attenuated as the beam propagates through the atmosphere and transfer of terrestrial wavelengths is dependent upon the vertical structure. In general the longwave radiation transfer leads to cooling whilst the absorption of solar radiation causes heating.

It is conceivable that the solar heating and terrestrial cooling could cancel at each point of the atmosphere; this is the state of radiative equilibrium. The structure of the radiative equilibrium atmosphere is of some interest for comparison to the real atmosphere. Various different types of zonally-symmetric equilibrium states of the middle atmosphere have been calculated; they are briefly discussed here.

Firstly, consider 'radiative-photochemical' equilibrium; here the temperature and radiatively active trace gas distributions are mutually adjusted until they are in photochemical as well as radiative equilibrium (e.g. Haigh, 1985). However, such a state has a very different ozone distribution to that observed; this occurs because of the long photochemical lifetime of ozone in the lower stratosphere, where the effect of transport becomes important. Comparison of the true atmosphere to the radiative-photochemical equilibrium state is not always the most useful comparison because of this large discrepancy in the lower stratosphere.

More appropriate estimates of the radiative equilibrium temperature structure are those obtained by using the observed (climatological) ozone distributions; the temperature structure of such states is calculated so that there is (as far as practicable) no net radiative heating at every latitude-height in the atmosphere. Two types of radiative equilibrium atmospheres have been calculated. The first, the state of radiative equilibrium, uses a fixed sun and constant ozone distribution for each month or season; the second allows the ozone distribution and solar insolation to progress through their annual cycles - the radiatively driven equilibrium (Fels, 1985; Shine, 1987) to obtain a time-dependent equilibrium structure. The latter balance has the advantage over the former in that latitudes at the edge of the polar night are not in perpetual darkness for the entire month. We shall denote the zonal-mean radiative equilibrium temperature field by $T_e(\phi,\eta,t)$.

If we assume thermal wind balance (equation (2.19)) the vertical structure of the zonal velocity at radiative equilibrium is given by:

$$\frac{\partial u_e}{\partial \eta} = \frac{c_p}{af}\frac{\partial T_e}{\partial \phi}.$$

The zonal-flow distribution may thus be obtained by specifying u_e at any height and integrating this equation using knowledge of the temperature distribution. Under the the same assumptions the northward flow, v_e, vanishes.

However, in the radiatively determined state the zonal-mean temperature distribution evolves; this implies that u_e is also time-dependent, so it cannot be strictly geostrophic: thus thermal wind balance is not strictly applicable. Equation (2.30) shows that the acceleration of the radiatively determined equilibrium zonal wind is given by:

$$\partial u_e/\partial t = -\Pi_{\phi e}$$

where the generalised potential vorticity flux in the radiatively determined state is given by

$$\Pi_{\phi e} = \sigma_{\eta e} v_e Q_e - [X].$$

Consequently, either a weak zonal dissipation or a weak northward velocity must exist if the zonal-mean radiatively determined state is to evolve. It is conceivable that a weak northward flow could result from the slow rearrangement of mass within isentropic layers, as suggested by the adiabatic version of the zonal-mean continuity equation (2.28):

$$\frac{1}{a\cos\phi}\frac{\partial}{\partial\phi}\left(\sigma_{\eta e}v_e\cos\phi\right) = -\frac{\partial\sigma_{\eta e}}{\partial t}$$

which may be integrated to obtain v_e from the evolution of the density field (assuming a boundary condition of no meridional flow at either pole). The argument of Tung (1982) suggests that the zonal-mean radiative equilibrium density field evolves very slowly, so that v_e must be small. Nevertheless, the existence of a weak, northward generalised potential vorticity flux is necessary for the radiatively determined state of the atmosphere to evolve.

2.3.2. The Forcing of the Atmosphere From Radiative Equilibrium.

In this section we consider possible forcing mechanisms to an atmosphere initially in its zonally-symmetric radiatively determined state; the objective of this conceptual experiment (adapted from ideas of Fels, 1985) is to investigate how departures from radiative equilibrium can occur. The slow evolution of the radiatively determined state is neglected (which is formally equivalent to introducing a WKB-type expansion on two timescales, assuming that the motions of interest occur much more rapidly than the evolution of the radiative balance so that they are separable; this *ad-hoc* assumption is not rigorously justified); this seems to be equivalent to using a radiative equilibrium, rather than a radiatively determined equilibrium state as the basis for the analysis. By this argument $\Pi_{\phi e}$ is important only on the long timescale and can be neglected for the analysis of the rapidly varying state.

Now express each quantity, s, as the sum of its radiative equilibrium distribution, $s_e(\phi,\eta,t)$, and its non-zonal deviation therefrom, $\Delta s(\lambda,\phi,\eta,t)$:

$$s = s_e + \Delta s.$$

In particular, since $v_e=w_e\equiv 0$:

$$(\langle v \rangle, \langle w \rangle) = [\sigma_\eta]^{-1}([\sigma_\eta \Delta v], [\sigma_\eta \Delta w]).$$

The evolution of the zonal-mean zonal velocity field can then be expressed (using equation (2.30)) as:

$$\partial([\Delta u])/\partial t = -[\Delta \Pi_\phi]. \tag{2.34}$$

Thus an anomalous (i.e. departure from the radiative equilibrium value) zonal flow requires the existence of an anomalous generalised potential vorticity flux. There are several mechanisms which are capable of inducing such anomalies in $[\Pi_\phi]$; these are now considered.

At radiative equilibrium there is no mean-meridional-circulation, so that $[\Delta \Pi_\phi]$ is non-zero only when there is a mean zonal force, [X], or if there are correlations between zonally-asymmetric eddies. The former requires the existence of either molecular viscosity or some small-scale (unresolvable) eddy forcing: we shall not treat this term here. The presence of large-scale eddies in the middle atmosphere is of some importance: we now briefly discuss their nature and the manner in which they arise.

Planetary-scale disturbances are excited in the troposphere by mechanical and thermal mechanisms. Such waves can propagate vertically in the atmosphere only if the zonal-mean flow is westerly and faster than the phase speed of the eddies (Charney and Drazin, 1961). Because the phase speed of a linear Rossby wave is given by $c=[u]-\beta k^{-2}$ (where $\beta=\partial f/\partial y$ and k is the zonal wavenumber) the longest waves have the slowest phase speeds, so only they can propagate to the middle atmosphere in the winter hemisphere. Shorter waves are trapped in the troposphere. The easterly mean flow in the summer hemisphere means that all planetary waves are trapped in the troposphere.

Thus, zonally-asymmetric departures from radiative equilibrium can occur in the winter hemisphere stratosphere: these can potentially cause a non-vanishing generalised potential vorticity flux as we now discuss.

Firstly, consider the possibility of thermal dissipation. A zonally-asymmetric disturbance causes longitudinal variations in the temperature and ozone structure leading to a non-zonal pattern of net radiative heating and cooling.

In an isentropic viewpoint the eddy vertical velocity no longer vanishes and there is the possibility of cross-isentropic transport of zonal momentum by the eddies; thus an anomalous generalised potential vorticity flux is induced. The problem of thermal dissipation is considered in chapter 4.

Secondly, consider the case when $[(\sigma_\eta v)'Q']$ does not vanish. This can occur if the eddies are mechanically dissipated or if they reach sufficient amplitude for finite amplitude effects to become important. The theory for the nonlinear behaviour of planetary waves as they approach a critical line (where the zonal flow and phase speed of the wave are equal) is discussed in Killworth and McIntyre (1985). The important part of the theory for our purposes is that as the wave approaches its critical line its amplitude grows so that nonlinear terms become important; material contours (e.g. distributions of potential vorticity on isentropic surfaces) are distorted so strongly that they buckle and become irreversibly deformed. Such wave-breaking was first isolated in distributions of potential vorticity calculated from stratospheric data by McIntyre and Palmer (1983, 1984). The redistribution of potential vorticity by these nonlinear perturbations results in a northward eddy component of the generalised potential vorticity flux by equation (2.34).

Thus, the zonal-mean flow can be accelerated from its radiative equilibrium state by the action of zonally-asymmetric perturbations (using relationship (2.34)). Concurrently, a temperature anomaly must be introduced: the atmosphere has been forced from radiative equilibrium. It is important to note that this departure from radiative equilibrium can arise entirely by the action of isentropic eddies (the latter case discussed). This may be the predominant mechanism in the middle stratosphere of the winter hemisphere; breaking planetary waves have been observed there by Al-Ajmi *et al.* (1985), Clough *et al.* (1985), Dunkerton and Delisi (1986), amongst others; the nonlinearity seems to act more rapidly than thermal dissipation at middle latitudes in the mid-stratosphere. In the tropics however, it is thought that radiatively dissipating equatorial waves are the dominant mechanism for accelerating the zonal-mean zonal flow.

2.3.3. The Observed Zonal-Mean State.

The zonal-mean structure of the observed stratosphere is not in radiative balance. Generally the winter polar stratosphere is much warmer than its radiative equilibrium temperature whilst the tropical stratosphere is slightly cooler than radiative equilibrium. The only region observed to be close to radiative equilibrium is near the summer polar stratopause. Thus, it is apparent that there is some force maintaining the atmosphere away from radiative equilibrium, acting against the 'radiative spring' (Fels, 1985) which attempts to relax the temperature structure back to its equilibrium value.

Just as in the idealised case discussed above it is the planetary waves which provide this forcing mechanism. Although the observed zonal-mean flow is not identical to its radiative equilibrium state it does show the same broad features: westerly in the winter hemisphere and easterly in the summer, so that the long planetary waves can still propagate vertically in the winter hemisphere but are evanescent in the summer hemisphere. Dickinson (1969) considered a simple model with a parametrisation of eddy forcing, chosen to represent that due to either thermally dissipated or weakly nonlinear perturbations; he found that the winter hemisphere of the middle atmosphere could be maintained away from its equilibrium state by the action of such eddies. Thus, perturbations forced in the troposphere and propagating vertically into the stratosphere (by the Charney-Drazin criterion, or some generalisation of it) are capable of maintaining the winter stratosphere from radiative equilibrium. This cannot be true of the summer stratosphere, so some other mechanism must be invoked (although it is not as far from radiative equilibrium as the winter hemisphere).

2.3.4. The Role of the Diabatic Circulation.

As in Tung (1982) partition the northward velocity into two components: $\langle v_D \rangle$ and $\langle v_T \rangle$, associated with diabatic and transient processes. The zonal-mean continuity equation (2.28) may then be separated into two parts:

$$\frac{1}{a\cos\phi}\frac{\partial}{\partial\phi}\left([\sigma_\eta]\langle v_D\rangle\cos\phi\right) = -\frac{\partial}{\partial\eta}\left([\sigma_\eta]\langle w\rangle\right) \tag{2.35}$$

$$\frac{1}{a\cos\phi}\frac{\partial}{\partial\phi}\left([\sigma_\eta]\langle v_T\rangle\cos\phi\right) = -\frac{\partial}{\partial t}\left([\sigma_\eta]\right) \tag{2.36}$$

It is clear that the diabatic northward velocity, on the left hand side of equation (2.35), is associated with a vertical gradient of cross-isentropic flow, which is a consequence of the atmosphere being maintained from radiative equilibrium by the eddy-forcing, as discussed previously. The direction of the northward diabatic flow is determined by the sign of the vertical gradient of vertical mass flux; this depends upon the density, which decreases with increasing entropy, and the vertical velocity, which depends on the net heating rate and hence the extent of departure of the atmosphere from radiative equilibrium. The transient component arises because of rearrangement of matter within isentropic layers leading to temporal changes in the zonal-mean density which requires a meridional transport of matter.

If a temporal average is performed and we assume the zonal-mean density field to be steady the transient northward flow, $\langle v_T \rangle$, defined by equation (2.36) must vanish. Then the mean-meridional-circulation is dominated by the diabatic component, $\langle v_D \rangle$.

Tung (1982) argued that this was also the case over long periods; he speculated that only during short periods of large-amplitude eddy activity, such as stratospheric sudden warmings, would the transient component be significant. This theoretical argument was substantiated by the observations of Gille *et al.* (1987), whose pressure-coordinate results showed that although the transient component was of considerable magnitude from day-to-day it was essentially reversible; its monthly-mean was negligible. It is clear that the diabatic circulation is the dominant part of the mean-meridional-circulation over long timescales; in Chapter 6 this will be applied to problems of zonally-averaged tracer transport.

The global scale of the mean-meridional-circulation is important. Equation (2.35) shows that even if there is a convergence of vertical flux at only one latitude a northward mean flow will be induced over a band of latitudes nearby. Consequently, at these latitudes $[\Pi_\phi]_M$ will become non-zero, so that acceleration of the zonal-mean flow will occur as a consequence of the redistribution of matter by the diabatic northward velocity. Thus, once the atmosphere is disturbed from radiative equilibrium at one latitude the induced mean-meridional-circulation can lead to a broadening of this region. Thus, if the winter hemisphere is maintained from equilibrium by the eddies, the

summer hemisphere can potentially be affected by the mean potential vorticity transport by the diabatic circulation.

Of course, this is not the only mechanism at work; the slowness of the relaxation of the summer hemisphere back to equilibrium after it moves out of winter must also be accounted for in a complete theory of the atmosphere.

2.4. The Three-Dimensional Diabatic Circulation.

2.4.1. The Importance of the Divergent Horizontal Motion.

The vorticity budget of the atmosphere is an essential aspect of the flow; in particular, the flux-form of the potential vorticity budget, equation (2.17), expresses that the local vorticity changes only in the presence of a divergent generalised potential vorticity flux, $\mathbf{\Pi}$.

Neglecting, for the moment, the diabatic and non-conservative components, we can write:

$$\mathbf{\Pi} = \sigma_\eta \mathbf{u} Q = \mathbf{U} Q = \mathbf{U}_R Q + \mathbf{U}_D Q, \tag{2.37}$$

where the horizontal mass flux, $\mathbf{U}$, has been partitioned into rotational and divergent components. In the normal manner a streamfunction and a potential can be introduced so that:

$$\mathbf{U}_R = \mathbf{k} \times \nabla_\eta \Psi \tag{2.38}$$

$$\mathbf{U}_D = \nabla_\eta \chi. \tag{2.39}$$

Then the divergence of the generalised potential vorticity flux may be expressed as:

$$\nabla_\eta \cdot \mathbf{\Pi} = \mathbf{k} \cdot \nabla_\eta \Psi \times \nabla_\eta Q + \nabla_\eta \cdot (Q \nabla_\eta \chi). \tag{2.40}$$

This represents advection of the potential vorticity by both the rotational and the divergent fluxes of matter and the product of the potential vorticity with the divergence of the horizontal flow.

Complete analysis of the vorticity budget therefore requires a knowledge of

the divergent flow field, since this is one of the forcing terms in the vorticity balance.

2.4.2. The Divergent Flow Field.

In isentropic coordinates there is an extremely simple expression for the horizontally divergent flux of matter: the continuity equation (2.6). From a knowledge of the diabatic heating rate at isentropic levels and the rate of change of density the continuity equation can be used as a 'diagnostic' relationship to obtain $\mathbf{U}_D$ (it is not strictly diagnostic since it requires a knowledge of the density change).

As in the zonal-mean case it is convenient to partition the divergent horizontal flow into diabatic and transient components, which are assumed to be independent. Then the continuity equation can be partitioned to give:

$$\nabla_\eta^2 \chi_D = -\partial W/\partial \eta \qquad (2.41)$$

$$\nabla_\eta^2 \chi_T = -\partial \sigma_\eta/\partial t, \qquad (2.42)$$

where $\mathbf{U}_D=\nabla_\eta\chi_D+\nabla_\eta\chi_T$ is the sum of the gradients of the potential fields for the diabatic (χ_D) and transient (χ_T) divergent flow.

The three-dimensional diabatic circulation may then be regarded as the motion associated with the cross-isentropic transport of matter, W, and the horizontal flow $\nabla_\eta\chi_D$, which together compose a three-dimensional non-divergent flow of diabatic mass transport (the divergence of the horizontal component of the flow is to compensate for any convergence in the vertical transport of mass). The transient component of the divergent flow is analogous to the transient northward component of the mean-meridional-circulation; it involves an isentropic rearrangement of matter within isentropic levels.

In steady flow the latter component vanishes, so that the divergent horizontal flow may be determined from knowledge of the diabatic heating rate: a consequence of the use of isentropic coordinates. Application to atmospheric data will be dependent upon the accuracy of the derived heating fields and, in the non-steady case, will depend upon the errors in determining the isentropic density; we shall consider this problem in chapter 7.

CHAPTER 3

THE DATA USED.

In this chapter the data used in the thesis are described. The primary data set is that of the Limb Infrared Monitor of the Stratosphere (LIMS) instrument, extended to the South pole with climatologies. In the first section the measurements and coverage of the LIMS experiment are outlined. Its validity is reviewed in section 2 and intercomparisons with independent data are made in section 3; thus any systematic biases, which will affect the derived fields, can be assessed. The climatological data are reviewed in section 4 and the manner of using this to create a global data set is described in section 5.

The results of chapters 5-7 will be presented in isentropic coordinates; thus the method of interpolation between log-pressure and entropy levels is described in section 6. Finally, section 7 contains a description of the method of estimating the wind and vorticity fields; an attempt is made to estimate the accuracy of these derived fields from the nominal accuracy of the temperature data.

3.1. LIMS Data: Coverage and Quantities Observed.

The LIMS instrument was mounted on the NIMBUS-7 satellite, launched on 24 October 1978. It was operational for seven months (25 October 1978 - 28 May 1979), its lifetime restricted by the mass of cryonogenic coolant it was allowed to carry. Here some of the orbital and experimental details affecting the coverage and quantity of the LIMS observations are mentioned; the review is not intended to be comprehensive: more details are to be found in Gille and Russell (1984) and references therein.

The platform was placed in a sun - synchronous orbit, with initial apogee and perigee altitudes of 953 and 940.5km, increasing to 959 and 945km respectively by the end of the experiment. The period of this slightly elliptic orbit was 103.98 minutes, increasing to 104.09 minutes as the altitude increased, with an equatorial crossing at local noon on the ascending portion of the orbit (node). Observations on the ascending node of the orbit were thus daytime measurements, whilst those on the descending node were nocturnal. The position and orientation of the LIMS instrument on the platform, with scanning out of the orbital plane to prevent direct illumination by the sun,

were such that the track of the tangent point was parallel to, and to the right of, the subsatellite point. Consequently, in the northern polar region the sensor viewed across the pole, whilst it scanned northward near the South pole; this geometry restricted the data coverage to between 64°S and 84°N, which is 86% of the globe. The orbital track for one day is shown in figure 3.1. Most of the measured signal is from a band 300km wide centred on the tangent point.

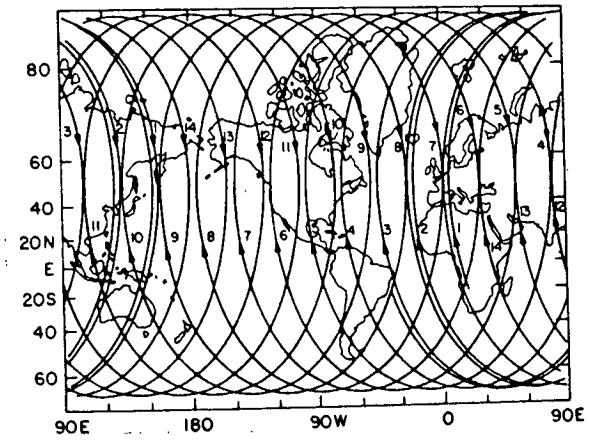


Figure 3.1. Mercator plot showing the tangent point for a 24 hour period, plus two additional orbits. Numbers indicate the orbital sequence; arrows pointing upward (downward) indicate ascending (descending) portions of the orbit. (From Gille and Russell, 1984.)

Radiance measurements were obtained in six channels, summarised in table 3.1. The wide and narrow channels centred on the 15μm CO_2 emission band were used to determine the temperature distribution as a function of pressure (CO_2 was assumed to be well mixed in the middle atmosphere). The temperature profile was then used with the radiance measurements in the remaining four channels to calculate the profiles of O_3, H_2O, NO_2 and HNO_3. The latter two constituents, present at mixing ratios of 10^{-9} by volume (ppbv), had not previously been measured by a satellite based sounder. Use of a limb sounding instrument enabled a good vertical resolution for these fields; the

field of view of each channel at the limb of the planet is also summarised in table 3.1. The methods and experimental details of the radiance measurements and and a summary of the inversion technique are summarised by Gille and Russell (1984), and in a series of validation papers (Gille *et al*, 1984a&b; Remsberg *et al*, 1984a; Russell *et al*, 1984a,b) and are not reviewed here.

Table 3.1. Summary of LIMS Channels.

Emitting Gas.	Wavelength Range(cm^{-1}).	Field of View at Limb (km). Vertical.	Horizontal.
CO_2(Wide)	579-775	1.8	18
CO_2(Narrow)	637-673	1.8	18
O_3	926-1141	1.8	18
H_2O	1370-1560	3.6	28
NO_2	1560-1630	3.6	28
HNO_3	844-917	1.8	18

It is important to appreciate the sensitivity of the retrieved gas profiles to the temperature. Because of the temperature dependence of radiative emission an underestimate of the temperature would lead to the deduction of there being too much tracer in the atmosphere; similarly an overestimate of temperature would result in the underestimation of tracer amount.

The individual data profiles have been objectively analysed to give globally mapped fields; it is these data which are used in this thesis. The processed data have a latitudinal resolution of 4° (note that the resolution of the radiance profiles at the limb is close to 1°, table 3.1) and a vertical resolution of around 3.5km; the standard pressure levels of the globally mapped data are summarised in table 3.2.

The global mapping was performed as follows. The individual profiles were screened and any bad data were removed; a Kalman filter was then applied, in which the data were treated as a time series at each latitude and height. This sequential estimation was performed both forwards and backwards in time for the zonal-mean and the Fourier coefficients for the first

six zonal waves. The average of the forward and backward time-marching was used as the field. The coefficients were stored on the Map Archival Tapes (MAT). In table 3.2 the fields stored on the MAT are summarised. In addition to the daytime and nighttime fields a diurnal average has also been calculated, except for NO_2 which shows a large diurnal variation. This average makes use of the interlacing of the ascending and descending nodes of the orbit in mid-latitudes (figure 3.1) to improve the zonal resolution to twelve Fourier coefficients.

Table 3.2. The Levels at Which Constituents are Mapped.

Pressure. (mbar)	Temp.	O_3.	H_2O.	NO_2.	HNO_3.
0.1	X,A,D	X,A,D			
0.2	X,A,D	X,A,D			
0.4	X,A,D	X,A,D			
0.5	X,A,D	X,A,D			
0.7	X,A,D	X,A,D			
1.0	X,A,D	X,A,D	X,A,D	D	
1.5	X,A,D	X,A,D	X,A,D	A,D	
2.0	X,A,D	X,A,D	X,A,D	A,D	X,A,D
3.0	X,A,D	X,A,D	X,A,D	A,D	X,A,D
5.0	X,A,D	X,A,D	X,A,D	A,D	X,A,D
7.0	X,A,D	X,A,D	X,A,D	A,D	X,A,D
10.0	X,A,D	X,A,D	X,A,D	A,D	X,A,D
16.0	X,A,D	X,A,D	X,A,D	A,D	X,A,D
30.0	X,A,D	X,A,D	X,A,D	A,D	X,A,D
50.0	X,A,D	X,A,D	X,A,D	A,D	X,A,D
70.0	X,A,D	X,A,D	X,A,D	A,D	X,A,D
100.0	X,A,D	X,A,D	X,A,D		X,A,D

X: Diurnal Mean Values; A: Ascending node (daytime) values; D: Descending node (nighttime) values.

From the objectively analysed temperature data the mean thickness of each of the standard pressure levels was calculated; these were combined with an

NMC analysis of the 50mbar height field to obtain the hydrostatically balanced height field at each pressure level. Dynamical fields have been derived from these height fields by several authors; a latitudinal streakiness was noted in these fields by Dunkerton and Delisi (1986), who attributed it to the Kalman filtering being undertaken independently at each latitude.

3.2. LIMS Data: Validity.

Because any errors in the LIMS temperature and trace gas measurements will affect the calculations of derived quantities, such as the radiative heating and cooling rates, it is important that we should appreciate the magnitude of these uncertainties. A series of validation experiments was performed on the LIMS data in an attempt to assess its accuracy and precision; this involved comparison of the satellite-derived profiles with *in-situ* measurements (rocket and radiosonde) and ground-based (radar) measurements and (as far as possible) repeated observations of the same column of atmosphere. The accuracy of the data (the difference between the mean of a large number of observations and the true value) and their precision (the standard deviation of the observations about their mean) could then be estimated; this is complicated by the lack of repeatability of observations from a travelling platform and by the lack of knowledge of the true atmospheric state at any time. Nevertheless, estimates of the accuracy and precision of the LIMS data have been provided.

A series of laboratory experiments was undertaken prior to the launch of the instrument so that the response of the sensors to various radiances could be determined and the effects of instrumental noise understood. This information was used in the retrieval of the atmospheric state from the radiance measurements. Comparison of the derived profiles with the independent measurements at a variety of locations around the orbit then enabled their validation. The experiments performed and the methods of analysis are summarised by Gille and Russell (1984) and presented in greater detail in the validation papers. A detailed review of these validation procedures is beyond the scope of this thesis; all we need to know are the validation statistics, which are summarised in table 3.3. Because we are using the synoptically smoothed MAT data it is most relevant to consider the data accuracy, rather than the precision, since small-scale, random errors are likely

to be smoothed out of the synoptic fields.

Temperature data are estimated to be accurate to ±2K in the stratosphere (p>1mbar) but at p<1mbar the LIMS measurements were persistently cooler than the rocketsonde measurements, whilst near 100mbar they were slightly warm. Ozone data were assigned an accuracy of 16% at 10mbar but this degraded to 40% at 100mbar and 0.5mbar; data below the peak mixing ratio at 10mbar are notoriously difficult to measure because the signal is swamped by that from the peak level; data in the low tropical stratosphere were contaminated by emission from convective cloud tops, so are unreliable. Water vapour data were designated to be 25% accurate between 100 and 1mbar.

Table 3.3. Accuracy and Precision of the LIMS Data.

Field.	Accuracy.	Precision.
T	<2K (p>1mbar)	<0.2-0.6K
O_3	16-41%	0.25ppmv
H_2O	18-36%	0.25ppmv
NO_2	20-50%	0.25ppbv
HNO_3	17-45%	0.15ppbv

3.3. LIMS Data: Intercomparison With Independent Measurements.

Here a brief review of the different characteristics of LIMS data compared to independent observations is given. It is unnecessary to go into great detail about the sources of data or to consider many days of data individually; the intention is to discuss how the LIMS data compares with other measurements, both in the averaged structure and the details of the flow.

Rodgers (1984) has compared the daily distributions of temperature and derived dynamical quantities for several days, using data from from various sources. Russell (1986) has compared monthly-mean temperature and trace gas fields obtained by several independent methods. The monthly-mean comparisons provide an assessment of any broad biases of any individual data set whilst the daily analyses can be used to estimate the structural detail in the different measurements. Comparisons of temperature, O_3 and H_2O are

reviewed here.

a. Temperature. Russell (1986) compared the monthly-mean temperature distributions of LIMS and SAMS. The zonal-mean cross sections generally agreed to within a few degrees Kelvin, both data sets showing the same basic structure. The horizontal distributions at 10mbar also agree well although LIMS data show slightly more structure than SAMS.

Daily temperature distributions were intercompared by Rodgers (1984) for six days, some dynamically active and some quiescent. As well as LIMS and SAMS data the objectively analysed Berlin and ECMWF fields (available in the lower stratosphere only) and the NMC analysis, which was compiled from *in-situ* observations and satellite data, were used. This latter field used the measurements of the Vertical Temperature Profile Radiometer (VTPR) until January 1979 when it was substituted by Stratospheric Sounding Unit (SSU) data. This sensor had a vertical resolution of around 10km throughout the stratosphere (Clough *et al*, 1985); however, one channel had a malfunction during the overlap with the LIMS experiment: thus its data are not accurate in the high stratosphere. The daily, zonal-mean cross sections showed that the three radiosonde analyses usually agreed to within 2K. LIMS and SAMS measurements generally agreed to within 2.5K poleward of 20°N and at pressures greater than 0.5mbar, although in the tropics there was slightly more disagreement (up to 6K) and at high levels (0.1mbar) the discrepancy could be as great as 20K. LIMS and NMC (VTPR) showed large differences (12K) on 26 January 1979 near 10mbar, 84°N (LIMS warmer) and 10K near 1mbar, 70°N (LIMS colder); a large amplitude, wavenumber-1 temperature wave was observed in the atmosphere at this time (eg Labitzke, 1981).

The daily perturbation fields also showed some differences, but these were usually small. In particular, SAMS and LIMS data tended to represent the first two zonal waves consistently and agreed well with the radiosonde analyses on the most disturbed days. NMC data agreed well with SAMS and LIMS at pressures greater than 5mbar, but above this level the agreement was less good, perhaps because of the retrieval mechanism used for the SSU data. More structure was evident in the LIMS data than the other fields, possibly a consequence of the better vertical resolution of LIMS.

The major conclusion is that LIMS temperature data agree well with the

other fields; they apparently have no gross bias, except for being rather cool above the stratopause. Further, the excellent vertical resolution of LIMS perhaps makes it the most suitable field for radiation transfer calculations, since the other data sets cannot represent small scale temperature structure.

b. Ozone. Russell compared LIMS ozone data to that of the Solar Backscattered UltraViolet Radiometer (SBUV) and the Stratospheric Aerosol and Gas Experiment (SAGE). Monthly, zonal-mean cross-sections were considered; three-dimensional distributions were also compared for the former pair of instruments.

In general, the SBUV and LIMS data were in closer agreement with each other than either of them was with SAGE. In January and February 1979 the maximum zonal-mean ozone mixing ratios observed by LIMS and SBUV occurred at the same location, but SBUV showed almost 2ppmv more ozone than did LIMS. In February LIMS detected more ozone at 10mbar, 10°N than SBUV; this could be due either to the LIMS temperatures being overestimated at high latitudes in February (because tight horizontal gradients are difficult to measure at high latitudes with a limb sounder) or to the poor vertical resolution of the SBUV data (8km, compared to the 3.5km of the processed LIMS data).

This latter feature is also apparent in the 10mbar ozone distribution in February, which also shows that LIMS detected more longitudinal structure in the ozone field than SBUV.

c. Water Vapour. Although no comprehensive, simultaneous data set was available for comparison, the LIMS water vapour data were compared to previous studies. Russell concluded that the LIMS data confirm the existence of a hygropause in the tropical stratosphere (after Remsberg *et al*, 1984b) and that the distribution supported the notion of downward transport at extra-tropical latitudes. Water vapour mixing ratios were found to increase with altitude above the low stratosphere at a rate consistent with that of methane oxidisation. Thus the LIMS observations are consistent with previous measurements.

3.4. Climatological Atmospheres.

Because LIMS data is not global it was necessary to extrapolate it, most notably to the South Pole; this enables application of a global constraint to the radiative energy balance of the stratosphere (eg Gille and Lyjak, 1986). Although we cannot hope to provide a perfect South polar region by extrapolation it should be possible to obtain a good approximation to the true atmosphere. It is fortunate that over most of the observation period it was the rather zonally symmetric summer pole which was not observed, since it would be far more difficult to extrapolate a highly disturbed field typical of the winter middle atmosphere.

The temperature climatology used was the proposed monthly, zonal-mean CIRA atmosphere (Barnett and Corney, 1985). It is a satellite-derived climatology of the atmosphere between pressure scale heights of 0.5 and 14.5 with a resolution of 0.5; the latitudinal resolution is 10°.

Two ozone climatologies have been used. The first, valid at pressures lower than 30mbar, is the CIRA climatology (Keating and Young, 1985); again this is a global, zonal-mean atmosphere, compiled from satellite data. The second, for pressures greater than 20mbar, is the monthly, zonal-mean Solar Backscattered Ultraviolet (SBUV) data set (McPeters *et al*, 1984). Some smoothing was necessary at the interface of these two climatologies: an equally weighted average of the data sets at 20mbar proved adequate, in that a smooth zonal field was attained. In some instances one or both of these distributions does not extend to the winter pole; on such occasions the polar-most value was retained toward the pole. Although crude, this method avoids imposing any artificial structure upon the climatological ozone distribution.

Climatological values of Carbon Dioxide mixing ratio (330 ppmv), Methane (0.88ppmv) and Nitrous Oxide (0.54ppmv) were used in the heating rate calculations. The latter two constituents play only a minor role in the heating, so a crude estimate of their mixing ratio should suffice. The Carbon Dioxide mixing ratio used here compares well with the amounts used by Gille and Lyjak (1986) and Kiehl and Solomon (1986) - 340 and 325 ppmv respectively.

For the terrestrial radiation transfer calculations a lower boundary radiance

was required. Accurate representation of this emission requires a knowledge of surface temperature and cloud top distribution. Ideally, daily observations of these quantities would be used, but even FGGE data does not include cloud observations. There is no comprehensive three-dimensional cloud climatology available; the closest to our requirements is that of Henderson-Sellers(1986), but this is only available for January and July, so is not really suitable. Consequently, the monthly, zonal-mean compilation of London (1957) was used.

3.5. Constructing a Global Data Set.

In this section the method of extrapolation of the LIMS data polewards and vertically is described. The poleward extrapolation is to enable the application of global constraints on the derived radiative heating rates. The vertical extrapolation is necessary for the longwave cooling calculation, which requires the specification of a troposphere (so that the radiation transfer within lower layers may be determined), and the ultraviolet heating calculation, which depends upon the incident solar radiation at each level, a function of the total amount of ozone in the column above this layer (as well as of the solar flux).

The poleward extrapolation of ozone and temperature fields was performed by matching the southernmost zonal-mean measurement to the gradient of the climatology, which ensured that a smooth field was obtained. Perturbations from the zonal mean state were forced to decay poleward as the cube of the cosine of latitude, so that the polar datum was the zonal-mean.

Vertical extrapolation was also performed by retaining the gradient of the climatology, matching it to the highest or lowest observation. This method causes any wave structure at the upper or lower level to be retained into the high atmosphere or the troposphere. Although unrealistic this was adequate, since neither the solar nor the terrestrial radiation transfer in the stratosphere are greatly influenced by the ozone and temperature distributions in the mesosphere. One would anticipate that tropospheric perturbations may not have a constant vertical structure, as implied by this extrapolation technique; however, tests described in the next section justify this somewhat crude approach. Additionally, Labitzke (1981) has shown that the wavenumber-1 perturbation in January 1979 had a deep, barotropic structure through the troposphere and stratosphere.

Water vapour distributions were extrapolated poleward simply by setting them equal to the southernmost observation; this is adequate, since water vapour plays only a minor role in the radiation budget of the atmosphere. A tropospheric water vapour mixing ratio was calculated from the temperature distribution using the method of Manabe and Wetherald (1967), which uses an analytic fit to determine the saturation vapour pressure from temperature and pressure for any tropospheric temperature. From the saturation vapour pressure a mixing ratio was obtained, assuming a relative humidity of 77% throughout the troposphere. The routine for the calculation of saturation vapour pressure was provided by Dr. R.S. Harwood.

3.6. Isentropic Interpolation of the Data.

Entropy levels selected for the study were η=6.2-7.8 at intervals of 0.1. The logarithm of potential temperature was used in preference to θ itself because of the more linear relationship with pressure scale height. The vertical resolution was selected to be comparable to that of the original LIMS data; it is actually slightly higher resolution, but this was felt adequate since it gives a smooth, regular interval.

The manner of interpolating the data from pressure to entropy levels is now described. Write the discrete entropy levels used as $\{\eta\}_k$, where $1<k<K$, and the exponents of the pressure scale-heights of the original data as $\{P\}_l$, where $1<l<L$ (figure 3.2). The objective of the interpolation is to interpolate data from the $\{P\}_l$ levels to the $\{\eta\}_k$ levels. Since the entropy at any scale height may be obtained from the temperature using the definition relating the entropy and scale height:

$$\eta = \log T + \kappa \log(P)$$

(where $\kappa=R/c_p$) it is straightforward to calculate $\eta(\{P\}_l)$. Interpolation between the $\{P\}_l$ and $\{\eta\}_k$ levels was then trivial: a second-order polynomial fit was used for chemical mixing ratios, radiative heating and cooling rates and temperature. Simple tests, not reported in detail, indicated that quadratic interpolation was adequate for our purposes, in that the smoothness of the fields was retained but spurious vertical structure was not introduced (this is a possible consequence of using high-order interpolation). Linear interpolation

resulted in a slightly 'jagged' vertical structure; cubic interpolation was similar to quadratic.

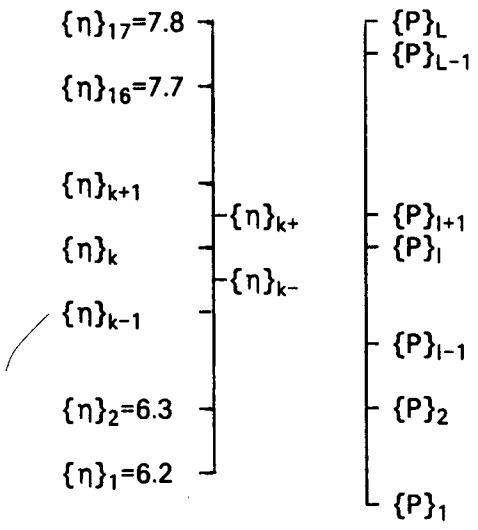


Figure 3.2 The entropy grid (left), showing the vertical distribution of entropy between k=1 and k=K=17, and its hypothetical relationship with the pressure scale-height grid (right), for $\{P\}_l$, $1<l<L$.

The density at each entropy level was obtained using a finite difference version of its definition (2.5):

$$\{\sigma_\eta\}_k = -(\{P\}_{k+}-\{P\}_{k-})/(g(\{\eta\}_{k+}-\{\eta\}_{k-}))$$

where k+ and k- denote the intermediate entropy levels. The pressure at these levels was again determined using quadratic interpolation from the LIMS data.

It was also necessary to construct the field of Montgomery Potential, $M=c_pT+gz$, at each isentropic level. Danielsen (1959) showed that simple addition of the geopotential and the specific enthalpy interpolated to isentropic levels results in an inconsistency in the hydrostatic relation leading to large errors in any subsequent isentropic analysis. He showed that if a tie-on level

is used, the Montgomery Potential being estimated from knowledge of temperature and geopotential height at one level, and the value at higher levels obtained by integrating the hydrostatic relation, equation (2.8), a consistent field was obtained. This method was used here. The Montgomery Potential at the lowest level was determined from the relation $\{M\}_1=c_p\{T\}_1+g\{z\}_1$, and that at higher levels obtained using:

$$\{M\}_k = \{M\}_{k-1} + (\{\eta\}_k-\{\eta\}_{k-1})c_p\{T\}_{k-}$$

where $\{T\}_{k-}$ is again obtained by second order interpolation, and is assumed to be a mean temperature for the layer between $\{\eta\}_k$ and $\{\eta\}_{k-1}$. The data at the lowest level were obtained using the NMC-based geopotential height and the LIMS temperature at 50mbar, provided on the LIMS data tapes. It is poignant to note that the LIMS geopotential heights were built up hydrostatically from this NMC height; any error in them would be worsened by additional interpolation to give M, whereas the method of Danielsen avoids this problem.

It is of considerable interest to determine how an error in the temperature field is manifested into the Montgomery Potential. The accuracy of the temperature is ±2K (table 3.3); if we assume that the temperature is 200K, this represents an error of 1%. Then the error induced in the Montgomery potential at each isentropic level is also 1% (using the finite difference method above). This error is accumulative, so that if the error in the temperature column was always (say) +2K, the error in the Montgomery potential at the k'th level would be k%. (This is neglecting any error in the base-heights used to obtain $\{M\}_1$). The impact of this on the dynamics will be discussed in the next section.

3.7. Calculation of the Vorticity and Velocity.

Because no direct observations of velocity and other dynamical quantities were obtained by the LIMS experiment it was necessary to derive such fields. The most widely used (and generally accepted) diagnostic relationship for deriving the velocity field is that of geostrophy (equation), which in centred difference form has the components:

$$\{u\}_{i,j} = - \frac{\{M\}_{i,j+1} - \{M\}_{i,j-1}}{2\Omega a^2 \sin\phi_j \cos\phi_j \, \Delta\lambda}$$

$$\{v\}_{i,j} = \frac{\{M\}_{i+1,j} - \{M\}_{i-1,j}}{2\Omega a^2 \sin\phi_j \Delta\phi}$$

where i and j represent the longitudinal and latitudinal indices (see Figure 3.3).

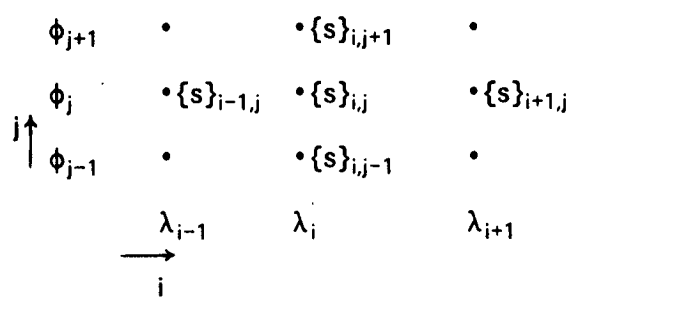


Figure 3.3 The horizontal grid (at constant k); latitude is represented by j and longitude by i. The finite difference notation for the quantity s is shown on the grid.

Recent studies have suggested that use of the geostrophic wind is unsuitable for some forms of diagnosis (Robinson, 1986; Boville, 1987). In particular, it transpires that the quasigeostrophic Eliassen-Palm flux vector exhibits regions of superfluous divergence at high latitudes, which do not correspond to the observation of the acceleration of the mean zonal velocity. Robinson has suggested that the use of the 'real' winds in place of the geostrophic values whilst retaining the quasigeostrophic formulation is sufficient to alleviate this problem (this study was based upon data generated by a primitive-equation numerical model). This approach seems somewhat inconsistent, because the Eliassen-Palm flux vector in its primitive form is not as simple as its quasigeostrophic equivalent (eg Andrews, 1987). The relationship between the northward eddy flux of potential vorticity and the divergence of the Eliassen-Palm flux is well known; Matsuno (1970) showed that energetic considerations required the isallobaric northward velocity to be used in estimates of [v'q']; this is most probably related to the problem

recently discovered with the eddy flux divergence.

There are alternative diagnostic systems which can be used to obtain 'higher order' wind fields than geostrophy; many of these are summarised by McWilliams and Gent (1980). An ageostrophic wind can be obtained by using the geostrophic wind in the absolute acceleration term of the momentum equation (this still neglects frictional terms and vertical advection). The linear and nonlinear balance equations, which are approximations to the divergence of the horizontal momentum equation in which the divergent component of velocity is neglected (and nonlinear terms are also ignored in the linear version) are a further set: solution requires the inversion of a second order differential operator, which in the nonlinear case has a rather complicated ellipticity condition (Houghton, 1968).

For the purposes of this study it was felt that the geostrophic approximation was adequate. There are several reasons for this. Firstly, the nature of the problems to be examined requires a knowledge of the broad features of the large scale flow, which the inclusion of non-geostrophic terms would not necessarily make more illuminating. Secondly, this study is not primarily concerned with estimating the acceleration of the mean-flow from eddy flux divergences. Thirdly, and perhaps most significantly, the lack of observed wind data with which to compare any diagnosed fields means that it is not possible to irrevocably prove that any of these other approximations are better than that of geostrophy: all are dependent upon the smallness of the Rossby number, so are least accurate in the same locations.

Smith and Bailey (1986) have performed a statistical analysis of the discrepancy between *in-situ* measurements of the wind with those derived geostrophically from LIMS data (at the standard pressure levels). Given the rather random uncertainties involved with the independent rocket soundings and the global uniformity of the LIMS measurements, a good agreement was obtained between the two data sets. One can imagine that the *in-situ* observations could be at their worst in regions of strong flow curvature and in fast moving jets, which is where the geostrophic approximation is also least good.

Of other recent studies of the middle atmosphere, geostrophic balance has been assumed by Butchart and Remsberg (1986), Dunkerton and Delisi (1986),

both using LIMS data, and Clough *et al.* (1985). McIntyre and Palmer (1983, 1984) used the gradient wind approximation in their precedent-setting study of the distribution of potential vorticity on the 850K isentropic surface.

Our calculations of vorticity have also assumed geostrophy; a centred difference form of the geostrophic vorticity (using the grid of figure 3.3) being:

$$2\Omega a^2 \sin\phi_j \{\zeta\}_{i,j} = \frac{\{M\}_{i+1,j} + \{M\}_{i-1,j} - 2\{M\}_{i,j}}{\cos^2\phi_j \Delta\lambda^2} + \frac{\{M\}_{i,j+1} + \{M\}_{i,j-1} - 2\{M\}_{i,j}}{\Delta\phi^2} - \frac{\{M\}_{i,j+1} - \{M\}_{i,j-1}}{\cos\phi_j \sin\phi_j \Delta\phi}$$

This assumption that the vorticity is well approximated by its geostrophic value is perhaps quite good, although it is most likely to break down in regions of of strong curvature.

A slight inconsistency may be introduced into the analysis by using the geostrophic relative vorticity to determine the potential vorticity, $Q = (\zeta+f)/\sigma_\eta$, but letting the density σ_η vary with horizontal position; this density in isentropic space is identical to the static stability in pressure coordinates, which quasigeostrophic theory does not allow to vary in the horizontal; there seems to be no complete formulation of quasigeostrophic motion in isentropic coordinates. However, if we consider that the transformation between pressure and isentropic space involves vertical displacements which are different at each horizontal position, a compatible formulation may need to allow σ to have a horizontal gradient in isentropic coordinates. Our data shows that the gradient is weak.

Because geostrophic theory is not formally applicable in the tropics, where the Coriolis acceleration no longer dominates the inertial term, the analysis is only performed poleward of 20°. The need to use centred difference formulae restricts our latitudinal coverage to 80°N.

At this stage it is convenient to consider the magnitude of errors induced in the dynamical fields by errors in the LIMS temperature profiles which affect

the Montgomery potentials. Recalling that the worst error in $\{M\}_k$ is $\pm k\%$, we must impose this error so that the worst possible effect is induced in the velocity and vorticity fields. The definition of velocity involves the difference between the Montgomery potential at two horizontal locations; thus, if each is in error by k%, the worst accumulative error at level k is 2k%; at η=6.7 this error is 12% whilst at at η=7.5 it is 28%. Consistent with this analysis of the errors in the velocity, the choice of profiles for determining errors in the vorticity is slightly restricted; it transpires that the worst possible error in $\{\zeta\}_{i,j,k}$ is 6k%. This has values of 36% at η=6.7 and 84% at η=7.5. These error bounds are very pessimistic since they assume entire temperature columns to be in error by ±2K, which is extremely unlikely; nevertheless, they reveal the amount of caution which must be exercised in interpreting the results of data analysis.

The last point of this chapter is to mention the smoothing of the data, which was performed to remove the 'streakiness' from the vorticity and velocity fields. A latitudinal average was performed to smooth out the two-gridpoint waves from the Montgomery potential fields before the vorticity and velocity were derived. The smoothed fields of Montgomery potential were almost unchanged by this smoothing (in that a polar stereographic plot of the same field before and after smoothing looked identical), but the vorticity field showed a much improved latitudinal structure.

CHAPTER 4

THE CALCULATION OF RADIATIVE HEATING AND COOLING RATES.

This chapter is concerned with the theory of radiation transfer in a plane parallel atmosphere and the application of this theory to atmospheric data; the first four sections review the theory, briefly describe the model used and examine zonal-mean cross-sections of the solar and terrestrial components of radiative heating rates. The difference between the zonal-mean of the three-dimensional fields and the zonal-mean fields calculated from the zonal-mean data is considered in section 5. Because the satellite data is subject to errors an attempt has been made to examine the sensitivity of the heating rate calculations to perturbations in the data structure; these experiments are reported in section 6. These errors are important, since they dictate the confidence which can be placed in the calculations of the components of heating and their sum; thus the validity of the results of chapters 5-8 is determined by the errors discussed in this chapter.

4.1. Introduction to Radiation Transfer.

In this section the equations governing the transfer of radiation in a plane parallel atmosphere will be introduced and the manner of obtaining the radiative heating rate from the radiative flux convergence will be outlined. The remainder of this chapter will build upon the concepts introduced here. The theory of radiation transfer is treated more comprehensively by Brasseur and Solomon (1986); both their treatment and a lecture course by Dr. K.J. Bignell at Imperial College in 1984 have influenced the presentation of the first four sections of this chapter.

4.1.1. Radiant Energy Transfer.

The radiance, I, is defined as the amount of energy, dE, traversing a surface, area dS, per unit time, dt, in solid angle $d\omega$ at angle θ to the surface:

$$I = dE/(dt\ dS\ d\omega\ \cos\theta) \tag{4.1}$$

(figure 4.1a). We may consider the atmosphere of Earth to be plane and parallel so that the surface dS is horizontal: θ is then the zenith angle; the azimuthal angle ϕ is defined in terms of the horizontal projection of the

direction of view, from some reference direction (figure 4.1b).

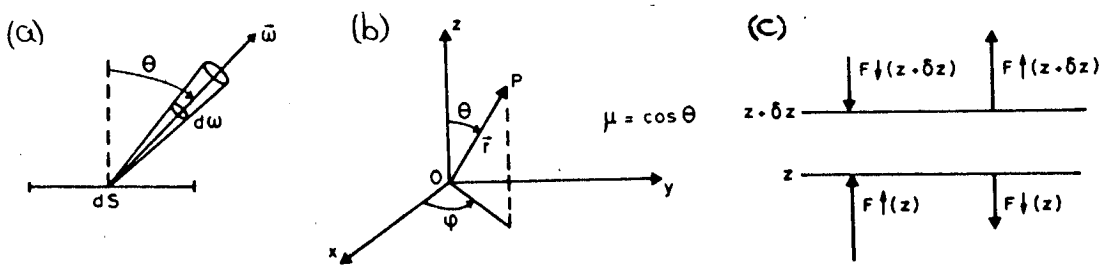


Figure 4.1. (a) Geometry of a pencil of light in solid angle dω traversing a unit of surface dS' at an angle θ to the normal; (b) definition of zenith (θ) and azimuthal (φ) angles; (c) balance of energy fluxes for a thin layer between z and z+dz in a plane parallel atmosphere. (From Brasseur and Solomon, 1986.)

To model the radiation transfer it is necessary to consider the monochromatic intensity of radiance:

$$I_\nu = dI/d\nu, \tag{4.2}$$

where ν represents the frequency. Integration of equation (4.2) over all frequencies gives the radiance, I, defined by equation (4.1). The radiative energy balance is expressed in Schwarzschild's equation:

$$\mu \frac{dI_\nu(z;\mu,\phi)}{dz} = -\rho_a k_\nu(z)\left[I_\nu(z;\mu,\phi) - J_\nu(z;\mu,\phi)\right] \tag{4.3}$$

where $\mu=\cos\theta$ is the cosine of the zenith angle. The source function, $J_\nu(z;\mu,\phi)$ represents the thermal emission of radiation of frequency ν per unit volume at height z in the direction (μ,ϕ); it also includes the radiation scattered in this direction but incident from other directions, although this will generally be neglected in this thesis. The extinction coefficient, k_ν, represents the amount of monochromatic radiance absorbed per unit of absorber which has density

$\rho_a(z)$.

For future reference the optical depth is introduced; this is defined as:

$$d\tau_\nu(z) = -\rho_a k_\nu(z)dz. \tag{4.4}$$

The passage of radiation through the atmosphere is easily understood in terms of optical depth: when τ_ν is small (the atmosphere is optically thin) a large part of the incident radiance is transmitted; an optically thick atmosphere absorbs a large part of the incident radiance; if the optical depth is infinite all of the incident radiation of frequency ν is absorbed in the atmosphere. Schwarzschild's equation (4.3) may be re-expressed in terms of optical depth:

$$\mu\frac{dI_\nu(z;\mu,\phi)}{\partial\tau} = I_\nu(z;\mu,\phi) - J_\nu(z;\mu,\phi); \tag{4.5}$$

this can sometimes provide a more illuminating analysis of the radiation transfer.

4.1.2. Flux Density and Radiative Heating.

The monochromatic irradiance, F_ν, often called the monochromatic flux density, is obtained by integrating the radiance over all zenith and azimuthal angles to a horizontal surface in a plane parallel atmosphere. Thus:

$$\begin{aligned} F_\nu(z) &= \int_0^{2\pi} d\phi \int_{-\pi/2}^{\pi/2} d\theta I_\nu(z;\theta,\phi)\cos\theta \\ &= \int_0^{2\pi} d\phi \int_{-1}^{1} d\mu I_\nu(z;\mu,\phi)\mu. \end{aligned} \tag{4.6}$$

The flux density may be partitioned into upward, $F\uparrow(z)$, and downward, $F\downarrow(z)$, components by integrating over upward and downward facing hemispheres, so that:

$$F_\nu(z) = F_\nu\uparrow(z) - F_\nu\downarrow(z), \tag{4.7}$$

where the upward (downward) flux is defined as the integral for $0<\mu<1$ ($-1<\mu<0$) in the definition (4.6).

A convergence of monochromatic flux density at any level implies that

there is a net input of radiative energy at that frequency, which causes a heating:

$$J_\nu(z) = -\frac{1}{\rho}\frac{dF_\nu}{dz} = g\frac{dF_\nu}{dp} \qquad (4.8)$$

where the latter part of this relationship depends upon hydrostatic balance. The total radiative heating rate, J, is obtained by summing this over all frequencies. In this thesis the quantity of most interest is the radiative heating rate in degrees per day, rather than the rate of change of energy, so the quantity J/c_p will be referred to as the radiative heating rate. The finite difference form of the definition of the radiative heating rate (equation (4.8)) for a layer of thickness $2\delta z$ centred at height z (figure 4.1c) may be written, using equation (4.7), as:

$$J_\nu(z) = -\left(\frac{F_\nu\uparrow(z+\delta z)-F_\nu\uparrow(z-\delta z)+F_\nu\downarrow(z+\delta z)-F_\nu\downarrow(z-\delta z)}{2\rho(z)\delta z}\right).$$

4.1.3. Black Body Emission.

The total flux emitted by a black body (also known as the exitance) is defined by the Stefan-Boltzmann law:

$$F = \sigma T^4 \qquad (4.9)$$

where the Stefan-Boltzmann constant $\sigma=5.67\times10^{-8}Wm^{-2}K^{-4}$. Because black body emission is isotropic the integral of equation (4.6) over the zenith and azimuthal angles gives:

$$F = \pi I. \qquad (4.10)$$

That is, the flux and the radiance are directly proportional. It is trivial to show that this relationship also holds for the monochromatic intensity and the upward and downward fluxes, provided we define the upward and downward radiances appropriately.

In accord with our previous discussion we are interested in the

monochromatic emission. For a black body the spectral distribution of emitted radiance is given by Planck's function:

$$B_\nu(T) = 2h\nu^3c^{-2}\left(\exp(h\nu/kT) - 1\right)^{-1} \qquad (4.11)$$

where Boltzmann's constant $k=1.38\times10^{-23}JK^{-1}$ and Planck's constant $h=6.63\times10^{-34}Js$; the velocity of light is $c=2.99\times10^{8}ms^{-1}$. Examples of the spectral distribution of radiant emission by black bodies are shown in figure 4.2; these Planck curves show that the radiance emitted by a black body is concentrated within a limited range of frequencies.

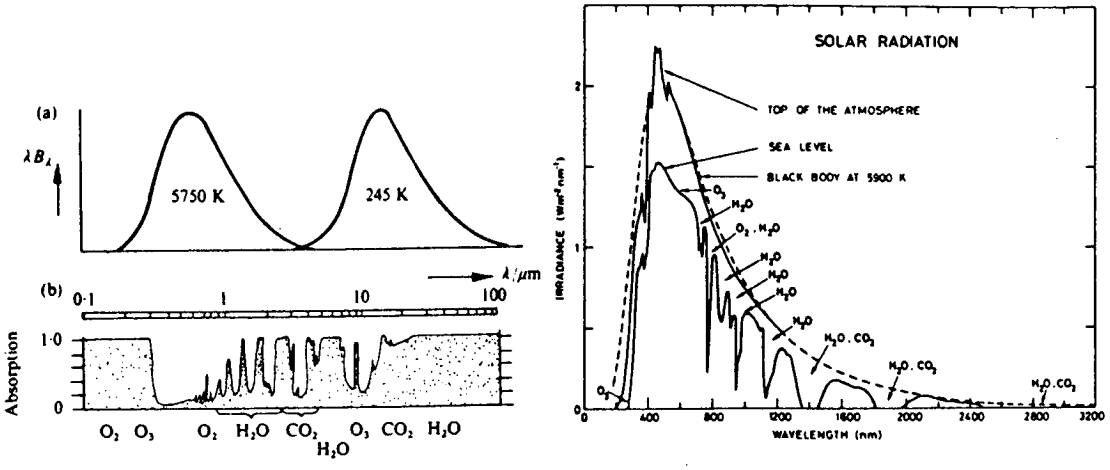


Figure 4.2. (a) Planck curves for black bodies at temperatures 5750K and 245K as a function of wavelength; (b) Total absorption of radiation in the atmosphere as a function of wavelength; (c) The solar irradiance incident at the top of the atmosphere and at sea level, with indications of the gases responsible for the absorption. (Parts a and b from Houghton, 1978; part c from Brasseur and Solomon, 1986.)

4.1.4. Separability of Solar and Terrestrial Radiation.

Assuming the Sun and Earth to be a black body emitters with temperatures of 5750K and 245K respectively, the spectra of emitted radiation have very little overlap (figure 4.2a). This is not a perfect assumption, since neither are

black bodies at any one temperature but the major point is that the great difference in temperature results in very little overlap between their emission spectra: consequently the solar and terrestrial components may be treated separately. This is the basis of most current models of radiation transfer in the atmosphere. In the solar part of the spectrum wavelengths shorter than 0.3μm (ultraviolet) are absorbed by O_3 and O_2 and are almost totally removed from the spectrum at the surface; in the visible part of the spectrum there is some absorption by O_3 and O_2 (figure 4.2b). This portion of the spectrum is considered in section 4.2. In section 4.3 the solar heating in the region 0.7–4.0μm is considered. Terrestrial radiation is considered in section 4.4; figure 4.2b shows that the most important gases for the terrestrial infrared radiation transfer are CO_2 (at 15μm), O_3 (at 9.6μm) and H_2O at a variety of wavelengths.

4.2. Solar Ultraviolet Radiation.

4.2.1. Theory.

Models of the transfer of solar radiation must consider the propagation of a beam through the atmosphere in a particular direction (except for the scattered component); thus it is convenient to modify the angular dependence in Schwarzschild's equation (4.3) and consider a beam incident at local zenith angle θ. Schwarzschild's equation thus reduces to:

$$dI_\nu(z;\theta) = -\rho_a k_\nu(z) I_\nu(z;\theta) \sec\theta dz.$$

Denoting the monochromatic radiance incident at the top of the atmosphere as $I_\nu(\infty)$ this equation may be integrated to give the Beer–Lambert law,

$$I_\nu(z;\theta) = I_\nu(\infty) T_\nu(z;\theta). \tag{4.12}$$

Here the transmission function, T_ν, denotes the fraction of the monochromatic radiance incident at the top of the atmosphere at angle θ which has been transmitted to the level z; mathematically it is defined by

$$\begin{aligned} T_\nu(z;\theta) &= \exp\left(-\int_z^\infty \rho_a k_\nu(z) \sec\theta dz\right) \\ &= \exp\left(-\tau_{\nu A}(z;\theta)\right), \end{aligned} \tag{4.13}$$

where $\tau_{\nu A}$ denotes the optical depth of the path between height z and the top of the atmosphere.

The zenith angle, θ, depends upon latitude, season and time of day, as discussed by Brasseur and Solomon (1986); furthermore, it is only for $\theta < 75°$ that the above formulation is adequate since at larger zenith angles the curvature of the atmosphere becomes important and the $\sec\theta$ factor describing the path length in equation (4.13) must be replaced by a Chapman function which includes the spherical geometry of the atmosphere (again see Brasseur and Solomon).

Equations (4.12) and (4.13) show that the radiance incident at any level z depends not only upon that incident at the top of the atmosphere but also upon the optical pathlength, which depends upon the amount of absorber in the path. Consequently, at any latitude any longitudinal variations in the absorber distribution above a particular level will cause zonal-asymmetry in the incident radiance at that level: this is known as the opacity effect; a consequence of it is that the absorption of solar radiation in the atmosphere can be greater at a longitude of small absorber amount than at a longitudinal maximum because of the extinction of the beam at higher levels. Thus, the maximum solar heating rate need not occur at the longitude-height of maximum absorber density. This effect will be important in the interpretation of the error analysis in section 4.6 and in discussions of the zonally-asymmetric radiative heating distribution in chapter 8.

4.2.2. Absorption by Ozone and Diatomic Oxygen.

Solar ultraviolet and visible radiation is absorbed by O_3 and O_2 in the middle atmosphere; in the troposphere the absorption by H_2O becomes significant, but its low concentration in the stratosphere means that this is negligible there. The quantum mechanical details of radiation transfer are discussed by Brasseur and Solomon (1986); here a brief summary of the spectral ranges of interest, the altitudes at which they are dominant and the magnitude of the heating rate caused by the absorption is given. (Also see table 4.1.)

a. O_3 absorption in the *Hartley Region* occurs primarily above 45km.

b. The *Huggins Bands* are important between 20 and 80km; again the absorption is by O_3.

c. The *Chappuis Bands* of ozone absorb very little radiation in the stratosphere, but play a significant role in the energy budget of the low stratosphere, where the total solar heating is small.

d. Both O_2 and O_3 absorb radiation in the *Herzberg Continuum* in the high stratosphere; the heating is around 10% as strong as that in the Hartley Region and Huggins Bands.

e. Absorption in the *Schumann-Runge Continuum and Bands* is unimportant below the stratopause. Similarly, absorption in the *Lyman-α lines* of O_2 occurs well above the region of interest of this thesis.

Table 4.1. A Summary of the bands of ozone and oxygen which absorb solar radiation and their altitudes of importance.

Name.	Spectral Range (nm).	Gas.	Altitude of Dominance.	Peak Heating [in tropics.]
Chappuis Bands	407.5-850	O_3	None	$\sim 1 Kday^{-1}$ near 30km.
Huggins Bands	277.5-360	O_3	20-80km.	$\sim 5 Kday^{-1}$, 40km.
Hartley Region.	242.5-277.5	O_3	>45km	$>10 Kday^{-1}$, 50km, $>2 Kday^{-1}$, 80km.
Herzberg Continuum.	206-242.5	O_3,O_2	35-55km	$>1 Kday^{-1}$, 50km.
Schumann - Runge Bands.	175-202.5	O_2	60-100km	$>10 Kday^{-1}$, 100km.
Schumann - Runge Continuum.	125-152	O_2	>80km	$\sim 10 Kday^{-1}$, 100km, Negligible below 80km.

This table was compiled from data given by Strobel (1978) and Brasseur and Solomon (1986).

4.2.3. The Model Used.

The model of solar ultraviolet heating used in this study was that formulated by Strobel (1978). It is an analytic parameterisation of the flux convergence in each of the intervals described in the previous section (table 4.1); it is an updated version of the original model of Lindzen and Will (1974). The code for the model was provided by Dr. J.D. Haigh and adapted for use in this study by Dr. R.S. Harwood.

The premise of the model is simple: assuming that the molecular cross-

sectional area in each of the bands is either constant or exponentially variable with frequency, data of the solar flux and molecular line-widths, with equations (4.12) and (4.13) integrated over the spectral range, were used to obtain an analytic fit for the effective cross-sectional areas in each band. It was necessary to partition the Huggins Bands into two spectral intervals to obtain an accurate formulation. As well as the direct solar radiation, Strobel included an approximate parameterisation of the absorption of scattered radiation in the Chappuis Bands, which is important to the energy budget in the lower stratosphere.

Good accuracy is obtained with this model: Strobel claimed that it gives results within 5% of those attained with line-by-line calculations of the solar heating rate. Rosenfield *et al* (1987) found that the largest errors are in the Huggins bands between 5 and 7mbar.

4.2.4. Review of Publications of Ultraviolet Heating Rate.

Zonal-mean cross sections of solar ultraviolet heating rates have been published in several past studies. Here, some of the more recent studies which used contemporary data sets and radiation transfer models are reviewed; wherever possible data for December are considered.

a. Gille and Lyjak (1986); LIMS data, 21 December 1978; figure 4.3a. The level of maximum heating occurs close to 1.5mbar (near 50km) between the South pole (where is is 14Kday^{-1}) and 20°N; it then slopes upwards through the middle latitudes of the northern hemisphere and is almost zero in the polar night. The heating rate at 10mbar (32km) at the South pole is around 3.5Kday^{-1}, whilst at 1mbar at the equator it is 10Kday^{-1}.

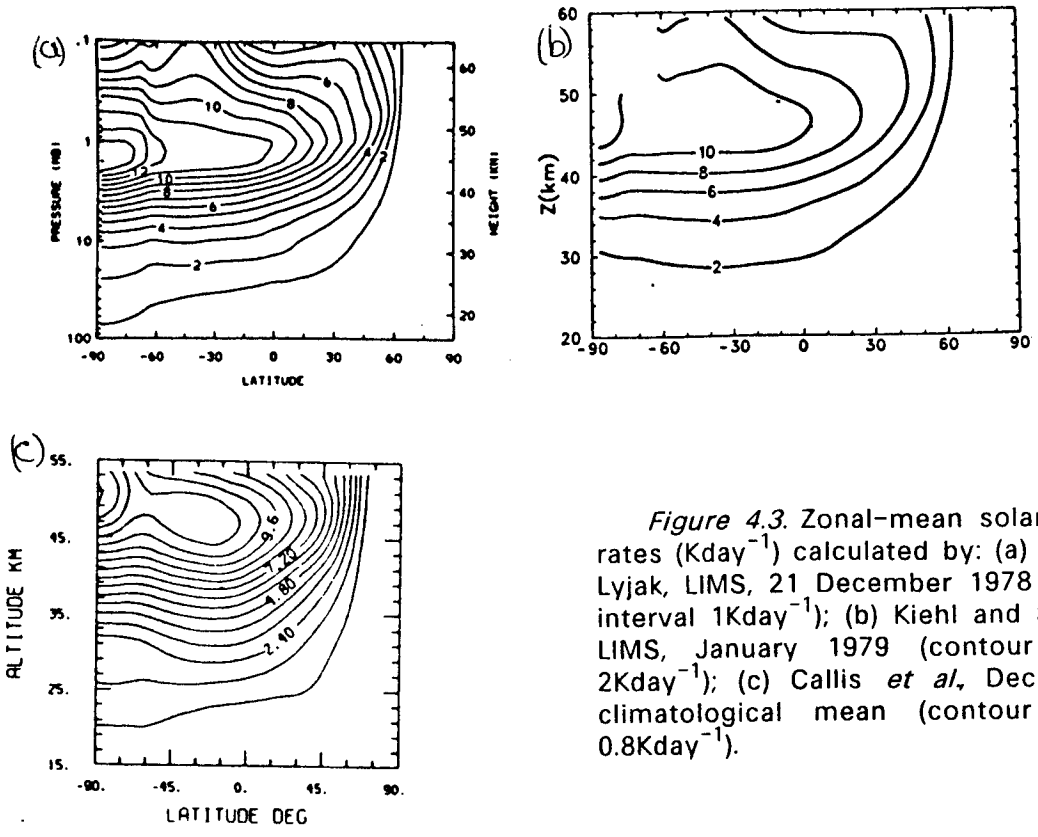


Figure 4.3. Zonal-mean solar heating rates ($Kday^{-1}$) calculated by: (a) Gille and Lyjak, LIMS, 21 December 1978 (contour interval $1Kday^{-1}$); (b) Kiehl and Solomon, LIMS, January 1979 (contour interval $2Kday^{-1}$); (c) Callis *et al*, Dec-Jan-Feb climatological mean (contour interval $0.8Kday^{-1}$).

b. Kiehl and Solomon (1986); LIMS data, January 1979; figure 4.3b. The level of heating maximum is very close to that of Gille and Lyjak although the heating is generally weaker; it is $11Kday^{-1}$ at the South pole. At 30km at the South pole there is $2Kday^{-1}$ heating rate, with $10Kday^{-1}$ at 45km at the equator. Agreement between these two studies is quite good at 45km, given that the data are from different dates; at lower levels (near 10mbar) Kiehl and Solomon calculate weaker heating than Gille and Lyjak.

c. Callis et al (1987); CIRA (1985) Climatology, December-January-February Mean; figure 4.3c. This seasonal-mean cross-section shows a level of heating maximum sloping downwards from the South pole ($13Kday^{-1}$ at 53km) to the equator ($10.4Kday^{-1}$ at 46km). The heating rate at 30km at the South pole is around $2.2Kday^{-1}$. These climatological results have their strongest heating rate at a higher level than the LIMS studies reviewed above; this could arise from the use of a seasonal average as well as use of a different data set.

4.2.5. Zonal-Mean Cross Sections of Ultraviolet Heating Rate.

The monthly, zonal-mean ozone and solar ultraviolet heating rates for December 1978 are shown in figure 4.4. The ozone distribution (figure 4.4a) has a maximum of almost 9ppmv near 30km, slightly to the South of the equator. At 45km the maximum mixing ratio occurs near 50°N, whilst at 50km it is near 60°N; it thus slopes upwards into the northern hemisphere. In the tropics the mixing ratio is smaller than 1.5ppmv at 100mbar.

The heating distribution shows a maximum of 17.3Kday^{-1} near 50km at the South pole, which is slightly stronger than the result of Gille and Lyjak (1986); this is not unexpected since different methods of interpolating the data to the pole and different radiative heating models were used. Our heating rate is considerably stronger than that obtained by Kiehl and Solomon (1986), which seems very weak in comparison to the other studies. In accord with the reviewed calculations the level of maximum heating in this study occurs at a pressure slightly lower than 1mbar over the entire southern hemisphere and the northern hemisphere tropics; the heating rate of 12.5Kday^{-1} in the tropics agrees well with that of Gille and Lyjak. At 35km at the South pole we calculate almost 3Kday^{-1} heating, which is slightly weaker than that of Gille and Lyjak; similarly, the heating at 10mbar in the tropics is slightly weaker than theirs.

Overall, the heating rates calculated in this study are slightly stronger at high levels (>40km) and slightly weaker in the low and middle stratosphere than in the other recent studies. Generally, our results are in better accord with those of Gille and Lyjak than those of Kiehl and Solomon (our heating rates for January, not shown here, are also stronger than those of Kiehl and Solomon).

4.3. Solar Infrared Radiation.

4.3.1. Introduction.

A slight contribution to the radiative balance of the middle atmosphere arises from the absorption of solar infrared radiation by water vapour, carbon dioxide, methane, nitrous oxide and diatomic oxygen. This heating is typically 10-15% of the magnitude of that caused by the absorption of solar ultraviolet

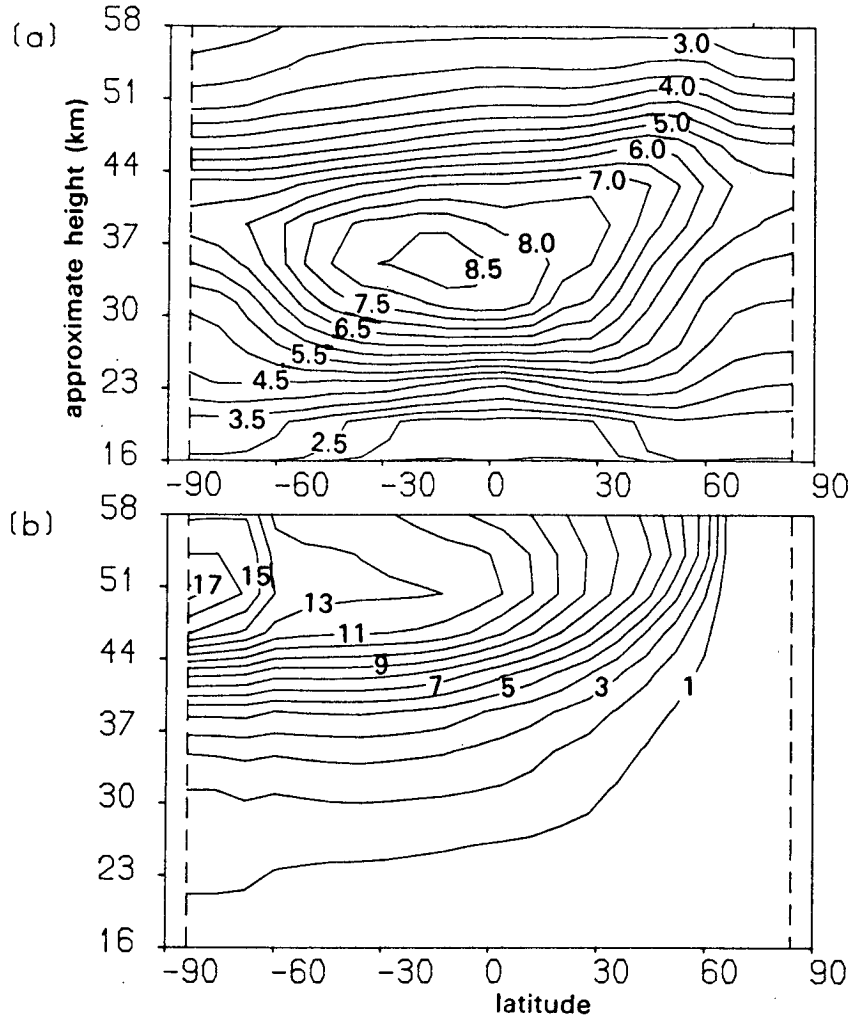


Figure 4.4 Zonal-mean cross-sections (a) LIMS ozone (ppmv, contour interval 0.5ppmv) and (b) solar heating rate ($Kday^{-1}$, contour interval [illegible] $Kday^{-1}$) for December 1978.

and visible radiation in the stratosphere (Haigh, 1984), but it plays an important role in the radiative balance, particularly in the low stratosphere.

The formulation of the radiation transfer of this part of the spectrum is identical to that for the ultraviolet radiation, since there are no sources in the atmosphere of Earth. A simple model of the heating at these wavelengths was used (Haigh, 1984).

The model has a simple flux-depletion formulation, following Rodgers (1967). As in Haigh (1984) the gases were assumed to have constant mixing ratios; values of 4ppmv, 0.88ppmv and 0.54ppmv were assigned to H_20, N_2O and CH_4 respectively. A more thorough treatment should allow the mixing ratios of all of the gases to vary; SAMS observations of CH_4 and N_2O show variations with latitude, height and time (eg Jones and Pyle, 1984); future studies should incorporate these variations and assess their importance. Because of the simplicity of this model, no tests of its sensitivity to changes in the data have been performed; note that even a 50% error in this heating calculation would represent only a 5% error in the total heating rate.

4.3.2. Zonal-Mean Cross Section.

The zonal-mean cross-section for December 1978 (figure 4.5) shows that the solar infrared heating rate is less than 10% as strong as the ultraviolet component at pressures lower than 5mbar. The pattern of heating is quite similar to that at shorter wavelengths; the strongest heating ($1.2Kday^{-1}$) occurs at the highest level (in the low mesosphere) at the South Pole. At higher pressures its importance increases somewhat, the heating rate of $0.1Kday^{-1}$ in the lower stratosphere is around 20% as strong as the ultraviolet and visible heating at the South Pole and of similar magnitude in the tropics. Thus, as the work of Haigh suggests, this component of the heating is most important to the radiative energy budget at low levels of the stratosphere where the net radiative heating is weak. Of the other recent studies, only Gille and Lyjak show the heating due to absorption of solar infrared radiation (by CO_2, H_2O and O_2); their heating rates were slightly weaker than those of this study at high levels but of comparable magnitude in the low stratosphere.

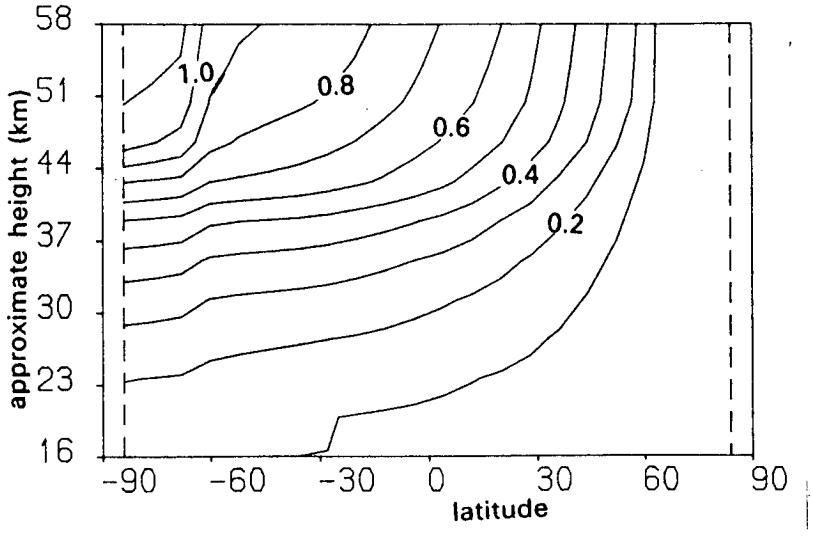


Figure 4.5. Monthly, zonal-mean solar infrared heating rate ($Kday^{-1}$) due to O_2, CO_2, H_2O, CH_4 and N_2O for December (contour interval $0.1Kday^{-1}$).

4.4. Terrestrial (Long Wave) Radiation Transfer.

4.4.1. Theory.

Calculation of the longwave radiation transfer within the atmosphere requires solution of Schwarzschild's equation (4.3) to obtain the flux convergence and hence the heating rate. In this section the essential details and approximations involved in its solution are outlined; again, a more thorough treatment is provided by Brasseur and Solomon (1986).

As already discussed, the majority of the energy transfer at terrestrial wavelengths is in the 15μm CO_2 band, the 9.6μm O_3 band and in various water vapour bands; minor contributions by the weak bands of these gases and by CH_4, N_2O, CO and various other gases are not treated in this study.

Solution of Schwarzschild's equation requires a knowledge of the source function. The Planck function, $B_\nu(T)$ (equation (4.11)) can be used as long as

the assumption of Local Thermodynamic Equilibrium (LTE) applies. This requires the energy levels in the model to be populated according to the Boltzman distribution, which occurs when they are dominated by collision processes; only then does the emission spectrum take the form of a Planck curve. The assumption of LTE applies in the stratosphere and lower mesosphere but breaks down at higher levels when the population of molecular energy levels starts to be influenced by radiative energy exchange. In the model used in this thesis LTE is assumed.

By analogy with the partitioning of the flux into upward and downward components (equation (4.7)) the monochromatic radiance may also be partitioned, so that Schwarzschild's equation (4.3) may be written:

$$\mu \frac{dI_\nu\uparrow(z;\mu)}{dz} = -\rho_a k_\nu(z)\left(I_\nu\uparrow(z;\mu) - B_\nu(T)\right)$$

and similarly for the downward radiance. Assuming that the ground emits as a black body at temperature T_S and that there is no downward radiance at the top of the atmosphere, the equations may be integrated to give the upward and downward radiance at any level of the atmosphere:

$$\begin{aligned} I_\nu\uparrow(z;\mu) &= B_\nu(T_S)T_\nu(z,0;\mu) \\ &\quad + \int_0^z B_\nu(T(z'))\{\partial T_\nu(z',z;\mu)/\partial z'\}dz' \\ I_\nu\downarrow(z;\mu) &= -\int_z^\infty B_\nu(T(z'))\{\partial T_\nu(z',z;\mu)/\partial z'\}dz'. \end{aligned} \tag{4.14}$$

The transmission function between any two levels z_1 and z_2 along the path is defined by:

$$T_\nu(z_1,z_2;\mu) = \exp\left(\int_{z_1}^{z_2} \partial\tau_\nu/\mu\right), \tag{4.15}$$

where the optical depth is as defined in equation (4.4). Solution of equation (4.14) with the definition (4.15) is, in principle, possible by considering each individual frequency of the spectrum; however, such a procedure is computationally expensive and too slow for use in most modelling and data analysis. Their solution is complicated further by the broadening of the spectral lines by collisions (often modelled as Lorentz broadening) and by the

motion of the individual molecules (Doppler). Below 40km the density is high enough for collision broadening to dominate, whilst Doppler broadening is dominant above 60km; at intermediate levels both processes are important so the line shape is determined by both processes: a Voigt distribution (which represents the convolution of Doppler and Lorentz profiles) is often assumed. A thorough description of these lineshapes and the manner of their application is beyond the scope of this thesis (see Brasseur and Solomon, and references within, for details); here, a brief summary of the approximations and method of solution is given.

Deriving the flux from the radiance involves integrating over the upward and downward hemispheres, as in equation (4.7). Rodgers and Walshaw (1966) showed that, to a very good approximation, this integration over zenith angle can be replaced by considering the passage of a collimated beam at the angle whose secant is 3/5 to the vertical (or equivalently, consider the vertical passage of a beam through an atmosphere 5/3 as deep as it really is). Defining the diffuse transmission function as $\mathcal{T}_\nu^*(z,z')=\mathcal{T}_\nu(z,z';\mu=3/5)$, the upward and downward fluxes are then defined by substituting F for I, $\mathcal{T}^*$ for $\mathcal{T}$ and πB for B in equation (4.14).

Rather than considering individual lines of the spectrum, intervals of lines are selected and mean transmission functions constructed for these spectral intervals. This involves the assumption that the lines are randomly distributed within each spectral interval and that the linestrengths have a particular distribution within each band (e.g. Goody, Malkmus - see Brasseur and Solomon for further details of these distributions). The mean flux density for each band can then be determined and the heating rate determined from its convergence.

4.4.2. The Model Used.

The longwave radiation transfer model used in this thesis was that of Haigh (1984). It was originally developed for use in a two-dimensional (zonally-averaged) numerical model of the middle atmosphere, so a computationally efficient parametrisation was necessary; this was achieved by partitioning the portion of the electromagnetic spectrum associated with terrestrial emission into five broad spectral intervals (table 4.2). These intervals are dominated by the 15μm band of CO_2 (although there is also a weak

contribution from water vapour in this interval), the 9.6μm O_3 band and various bands of water vapour.

Table 4.2. The Spectral Intervals of the Longwave Radiation Transfer Model.

Spectral Interval (cm^{-1}).	Contributing Gases.
0-555	H_2O (rotation)
555-775	CO_2 (vibration-rotation)
	H_2O (rotation)
775-960	H_2O (rotation)
960-1080	O_3 (vibration-rotation)
1080-2200	H_2O (vibration rotation)

The effective transmittances for the CO_2 and O_3 bands were derived by Haigh from the data of Ramanathan (1976). H_2O transmittances were derived from spectroscopic data. The correction in the second interval for the overlap of H_2O with CO_2 bands was treated by assuming random overlap and regarding the H_2O contribution as a random correction to the CO_2 absorption and emission (this is a very small correction).

4.4.3. Review of Published Cross-Sections.

a. Gille and Lyjak (1986); LIMS data, 21 December 1978; figure 4.6. The radiative cooling rates due to CO_2, O_3 and H_2O are treated separately. The strongest cooling due to CO_2 is $11Kday^{-1}$ near 1mbar at the South Pole; the altitude of strongest cooling follows the stratopause, remaining at a pressure of slightly greater than 1mbar except in high latitudes of the northern hemisphere where it slopes upwards to 0.5mbar. Cooling in the 9.6μm band of O_3 also peaks at the stratopause, exceeding $3Kday^{-1}$ at the South pole. In the low stratosphere the radiation transfer in this band results in a slight heating, exceeding $0.5Kday^{-1}$ in some parts of the tropics; Gille and Lyjak noted that the distribution of this heating was sensitive to the distribution of cloud in the troposphere and that the heating maxima are at latitudes of

minimum convective cloud cover; here a greater proportion of the radiation emitted by the surface reaches the low stratosphere, leading to stronger absorption. The water vapour cooling was generally much weaker than that due to the other two gases, peaking at 1.5Kday^{-1} at the summer polar stratopause; however, in the low stratosphere, where the CO_2 cooling is less than 1Kday^{-1}, water vapour contributed a further 0.25Kday^{-1} to the cooling, which represents a significant proportion of the total radiative energy budget.

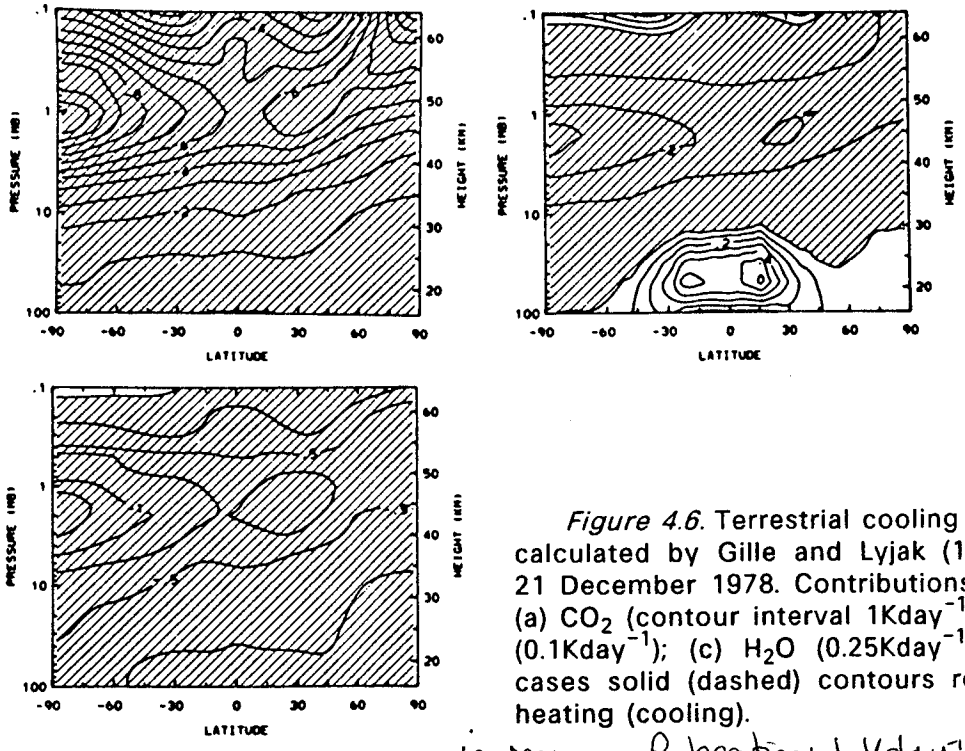


Figure 4.6. Terrestrial cooling (Kday^{-1}) calculated by Gille and Lyjak (1986) for 21 December 1978. Contributions due to (a) CO_2 (contour interval 1Kday^{-1}); (b) O_3 (0.1Kday^{-1}); (c) H_2O (0.25Kday^{-1}); in all cases solid (dashed) contours represent heating (cooling).

in regions of heating; 1 Kday^{-1} in regions of cooling

b. Kiehl and Solomon (1986); LIMS data, January 1979; figure 4.7. At the summer stratopause the cooling rates were 10Kday^{-1} for CO_2, 3Kday^{-1} for ozone and 1Kday^{-1} for H_2O; all of these are slightly weaker than the results of Gille and Lyjak, although since they are from a different month this cannot be interpreted definitively. Since no clouds were used in these calculations there was less structure in the distribution of O_3 heating in the low tropical stratosphere; cooling is calculated in this band at northern hemisphere mid-

latitudes where Gille and Lyjak calculate a slight heating. Lack of contours on the H_2O cooling cross-section makes interpretation of the structure of this field difficult.

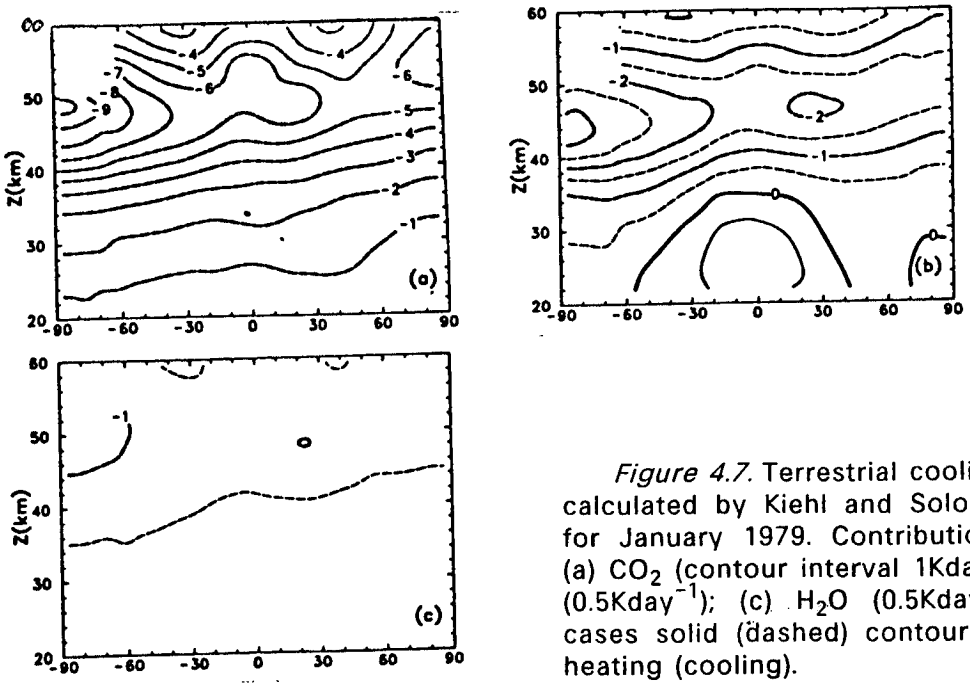


Figure 4.7. Terrestrial cooling ($Kday^{-1}$) calculated by Kiehl and Solomon (1986) for January 1979. Contributions due to (a) CO_2 (contour interval $1Kday^{-1}$); (b) O_3 ($0.5Kday^{-1}$); (c) H_2O ($0.5Kday^{-1}$); in all cases solid (dashed) contours represent heating (cooling).

c. Callis et al (1987); CIRA Climatology, Dec-Jan-Feb Mean; figure 4.8. The seasonal averaging performed in this study is the most probable explanation of the cooling rates being weaker than the others. The cooling rates at the South polar stratopause are $9Kday^{-1}$ (CO_2), $2.6Kday^{-1}$ (O_3) and $1Kday^{-1}$ (H_2O); at all latitudes the peak cooling occurred at the stratopause. The ozone heating exceeded $0.5Kday^{-1}$ in the equatorial low stratosphere; there is no structure in this field because a homogeneous cloud field was assumed, with a fractional cover of 0.446 at 5km altitude.

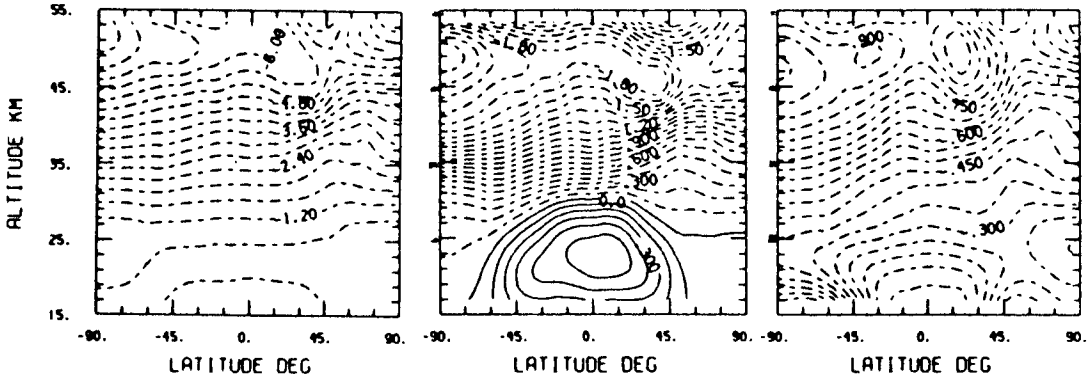


Figure 4.8. Terrestrial cooling ($Kday^{-1}$) calculated by Callis *et al.* for climatological Dec-Jan-Feb. Contributions due to (a) CO_2 (contour interval $0.8Kday^{-1}$); (b) O_3 ($0.1Kday^{-1}$); (c) H_2O ($0.05Kday^{-1}$); in all cases solid (dashed) contours represent heating (cooling).

d. Overview. In all cases the peak cooling occurs at the South polar stratopause; typically the CO_2 cooling accounts for around 75% of the total, with O_3 and H_2O playing lesser, but important roles. At lower levels, however, the balance is changed: H_2O contributes as much as 50% as CO_2 to the total cooling, whilst ozone causes a net heating (note that Gille and Lyjak calculate a very slight heating in the CO_2 band at 100mbar).

These results illustrate the delicacy of the radiative balance in the lower stratosphere; there is slight warming in the infrared and ultraviolet bands of ozone and slight cooling by CO_2 and H_2O, so that the net heating rate is a small residual of these components. It is in this region that the heating rates are most sensitive to the tropospheric state (as we shall see later), so we must interpret results for this part of the atmosphere with caution.

4.4.4. Zonal-Mean Terrestrial Heating Rates.

Zonal-mean heating and cooling rates for December 1978 are described here; the components due to CO_2, O_3 and H_2O are considered in turn (although recall that the CO_2 cooling also contains a weak contribution from H_2O - table 4.2).

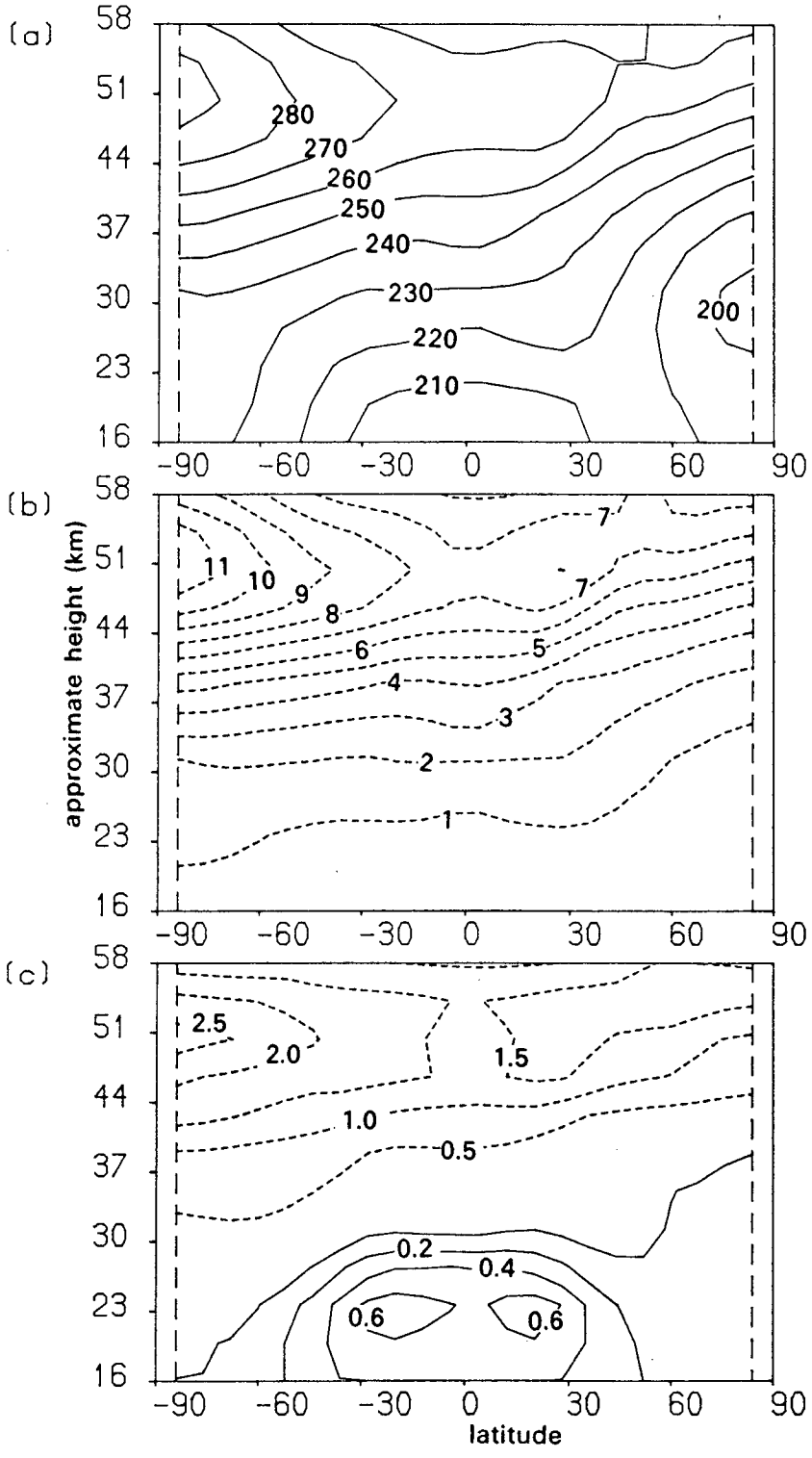


Figure 4.9. Zonal-mean cross-sections for December 1978 of (a) LIMS temperature (K, contour interval 10K); (b) cooling due to the 15μm bands of CO_2 ($Kday^{-1}$, contour interval $1Kday^{-1}$); (c) cooling due to the 9.6μm bands of O_3 ($Kday^{-1}$, contour interval 0.2 in regions of heating - solid line, and 0.5 in regions of cooling - dashed line). (Figure continues.)

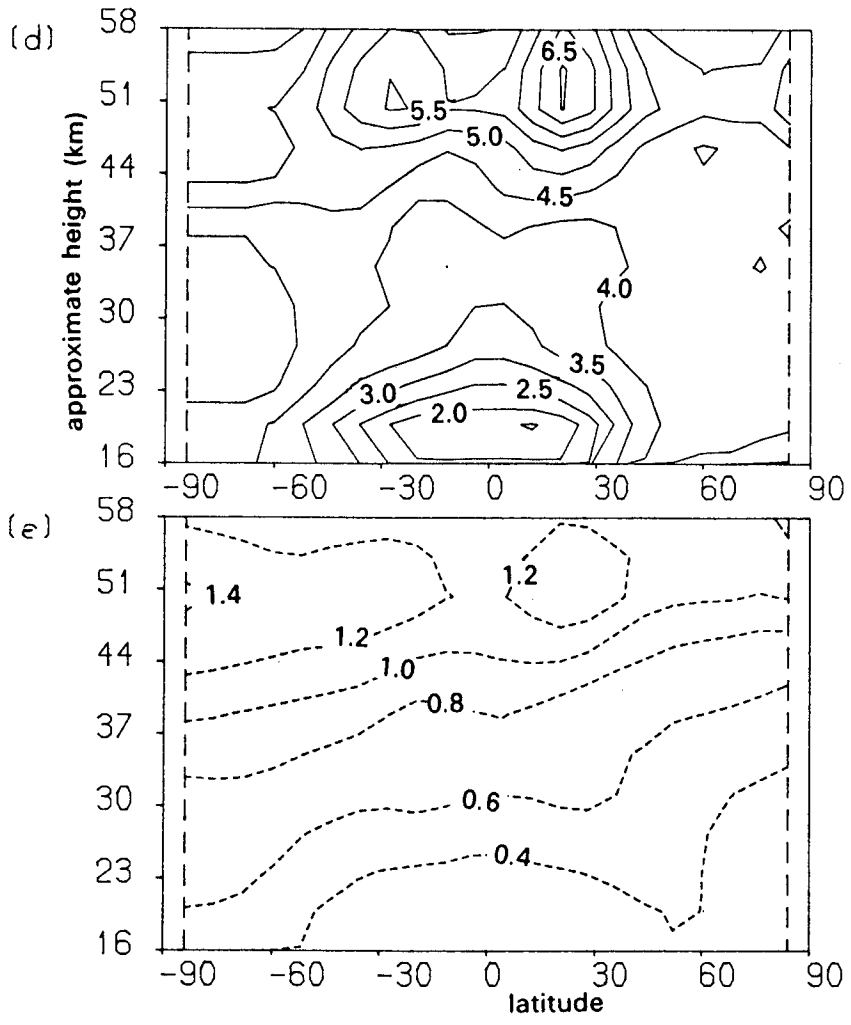


Figure 4.9 (continued). (d) LIMS water vapour (ppmv, contour interval 0.5ppmv) and (e) H_2O cooling ($Kday^{-1}$, contour interval $0.5Kday^{-1}$).

a. Carbon Dioxide, figure 4.9b. The strongest cooling of 11.9Kday^{-1} occurs at the South polar stratopause (~50km) where the temperature is 294K (figure 4.9a); this cooling rate agrees well with that of Gille and Lyjak, so is rather stronger than those of Callis *at al.* and Kiehl and Solomon. At the tropical stratopause our cooling rate is 7.4Kday^{-1} whilst at the North Pole, where the stratopause is rather higher, it is 8.4Kday^{-1}; these are both slightly stronger than all three reviewed studies.

In the middle stratosphere (near 30km) CO_2 causes cooling of around 2Kday^{-1} in the southern hemisphere and tropics; this decreases to slightly less than 1Kday^{-1} in the North Polar region. This agrees quite well with both of the other LIMS calculations, but Callis *et al.* find much stronger cooling near the North Pole: this must be attributed to their use of a seasonal climatology which includes the signatures of sudden warmings in January and February.

b. Ozone, figure 4.9c. In accord with previous results the terrestrial radiation transfer by ozone leads to heating in the low stratosphere, particularly in the tropics. The inclusion of cloud in our calculations has resulted in a twin maximum of heating (0.7Kday^{-1}) in the tropics; this is slightly stronger and further from the equator than the results of Gille and Lyjak, but reaffirms their conclusion that it is important to include cloud in calculations of the radiation transfer in the 9.6μm band of ozone.

At the South Polar stratopause ozone cools by 2.7Kday^{-1}; in the tropics and northern hemisphere this is reduced to 1.5Kday^{-1}; this is in good, although not perfect, agreement with the results of both Gille and Lyjak and Kiehl and Solomon but the seasonally averaged calculations of Callis *et al.* are rather different.

c. Water Vapour, figure 4.9e. In the high stratosphere H_2O cools by >1Kday^{-1} at all latitudes, with a maximum cooling of 1.4Kday^{-1} just below the stratopause at the South Pole. This represents a contribution of 10% to the total cooling rate; however, in the low stratosphere (near 50mbar) H_2O cooling is around 0.5Kday^{-1}, which is comparable to the CO_2 cooling and the heating at this level in the 9.6μm band of O_3. Water vapour makes an essential contribution to the radiative balance of the low stratosphere.

4.5. Linearity of the Zonal-Mean Heating Rates.

4.5.1. The Problem to be Addressed.

The radiative heating and cooling rates discussed in the previous three sections were calculated using zonally-averaged fields of temperature, ozone and water vapour. It is of some interest to determine how linear these calculations are: that is, does the zonal mean of the three-dimensional heating rate, $[J(T,O_3)]$, differ substantially from that calculated using zonal-mean fields, $[J([T],[O_3])]$. This topic is addressed in this section.

4.5.2. A Simple Experiment.

Using the three-dimensional LIMS temperature and ozone data a simple experiment has been performed in an attempt to quantify the size of the difference between the two calculations as a function of wave amplitude. Using the zonal-mean data and the observed wavenumber-1 for 25 January 1979 (when a large-amplitude wavenumber-1 was observed in the troposphere and stratosphere, Labitzke, 1981; see figure 4.10) the heating and cooling rates have been calculated at 60°N, and the zonal-mean of the three-dimensional calculation compared to the calculation using zonal-mean data only. In order to examine the response of this calculation to varying wave amplitudes, the temperature and ozone waves have been scaled until the peak of the temperature wave reached 50K; the amplitude of the temperature wave at levels away from the peak was scaled by the same ratio, as was the ozone wave at all levels. The relative phase between the two waves was also retained.

The results of this analysis showed that the difference in the net heating rate never exceeded 10% for a temperature wave with maximum amplitude 25K; this occurred at the level of maximum wave amplitude in the middle stratosphere. At this level the difference in the solar ultraviolet heating rate was less than 1%; the peak difference for the solar heating was near the stratopause (around 5%) where the solar irradiance is strongest: it is fortunate that the large amplitude waves occur where the solar illumination is weak. At 25K amplitude, the CO_2 cooling showed a difference of $<6\%$ near 30km and $<1\%$ above 40km, O_3 cooling showed rather stronger differences, ranging from 5-10% in the high stratosphere, with anomalously large differences at the

level of change from heating in the low stratosphere to cooling in the high stratosphere; the total longwave cooling changed by less than 2% at high levels (above 40km) and by less than 7% in the rest of the stratosphere.

As the wave was allowed to grow the differences became more significant, doubling when the peak amplitude of the temperature wave was 35K and being more than four times as strong when it reached 50K. However, it is unusual to observe waves with amplitude larger than 25-30K in the middle atmosphere, so such results are probably insignificant. The errors incurred by using zonal-mean fields in the calculation of zonal-mean heating fields are probably acceptable given the massive increase in computational resources required to calculate three-dimensional fields and zonally-average them.

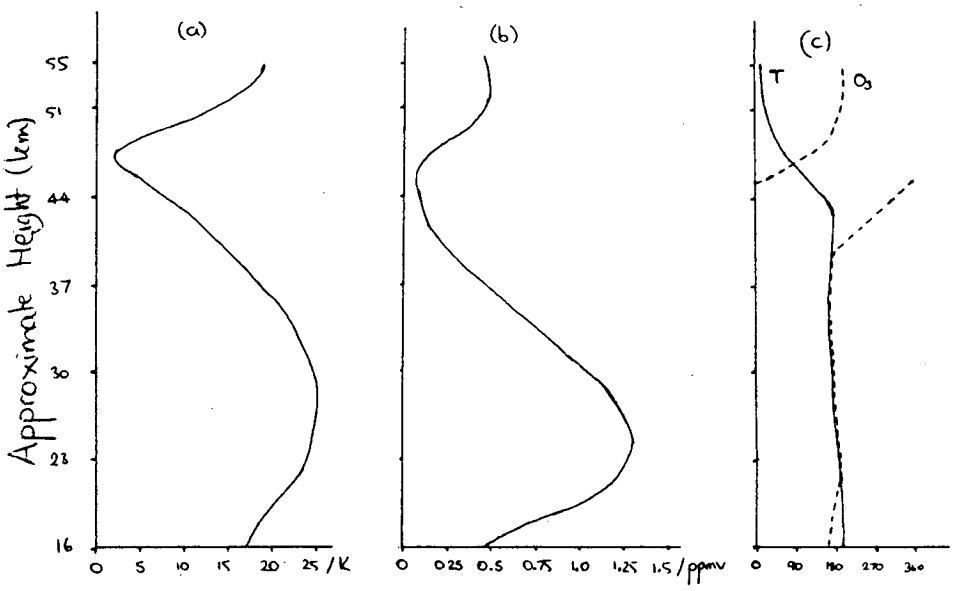


Figure 4.10. Amplitudes of (a) Temperature (K) and (b) Ozone (ppmv) wavenumber 1 and (c) their phases as a function of height on 25 January 1979.

4.6. Analysis of the Sensitivity of the Calculations to Data Errors.

4.6.1. Introduction: Motivation and Methods.

In an attempt to place bounds upon the accuracy of the heating and cooling rates calculated from LIMS data a series of perturbation experiments has been performed in which the data were perturbed and the response of the heating and cooling rates examined. There are two levels at which this error analysis must be performed. For this chapter and chapter 8 an estimate of the error in the solar heating rate and the components of cooling rate must be obtained; for chapters 5-7 an estimate of the accuracy of the net heating rate is needed; both are considered in this section.

Errors in the calculations arise from three areas. Firstly, the ozone distribution above the stratopause affects the incident solar radiation in the stratosphere; however, there are relatively small quantities of ozone above the stratopause than in the stratosphere, so it transpires that the stratospheric heating rates are quite insensitive to changes in the upper atmospheric ozone distribution. No significant longwave radiation is emitted downwards by the high atmosphere.

Secondly, the tropospheric structure determines the amount of upward radiation, so that variations in tropospheric temperature, ozone and water vapour distributions affect the upward flux at the tropopause. The response of the stratospheric heating rates to variations in the tropospheric state are discussed in section 4.6.3. Thirdly, errors in the LIMS data themselves affect the calculated emission and absorption of radiative energy in the stratosphere; this is discussed in section 4.6.3.

The approach adopted for these perturbation analyses was to calculate the heating and cooling rates for a variety of atmospheric columns, including highly disturbed winter mid-latitude profiles, tropical columns and summer hemisphere columns. These data columns were then perturbed and the heating rates recalculated; the fractional change in the heating and cooling rates was then determined. For the tropospheric tests the entire tropospheric data column was perturbed. Stratospheric tests allowed either the entire column or the data at one level to vary. In all cases the response of the heating rate to perturbations in only one quantity was examined, but an additional test

allowed the LIMS temperature to increase whilst the ozone decreased (and *vice versa*). Obviously, in the real atmosphere all quantities should vary together; for the purposes of this thesis it was felt adequate to allow the quantities to vary individually, so the response to errors in each field could be examined. The range of tests performed is not, by any means, comprehensive – there are a large number of possible variations which could be included; the tests to be reported were performed to attempt to place absolute bounds upon the errors in the data, given the physical conditions by which the satellite profiles were obtained.

Although the largest errors in the data can be expected on single days when there was highly disturbed flow (such as late January 1979 in the northern hemisphere) the errors calculated for such profiles were qualitatively similar to those for other profiles. This is most true of the radiation transfer at terrestrial wavelengths, since the solar component has a stronger latitudinal variation. It is convenient to consider the errors for one profile for one day; the profile chosen for discussion is that from 60°E, 60°N on 25 January 1979. Before discussing the perturbation tests, the heating and cooling profiles for this day are discussed.

Table 4.3. The vertical distribution of heating and cooling rates ($Kday^{-1}$) for 60°E, 60°N on 25 January 1979.

Approximate Height (km)	J_{SOL}	J_{NET}	J_{LW}	J_{CO2}	J_{O3}	J_{H2O}
55.0	4.1	-9.6	-14.0	-10.9	-1.8	-1.3
51.5	4.8	-9.0	-14.1	-10.2	-2.5	-1.4
48.0	3.6	-6.0	-9.9	-7.2	-1.5	-1.1
44.5	1.9	-5.2	-7.2	-5.1	-1.2	-0.9
41.0	1.1	-4.3	-5.5	-4.1	-0.5	-0.9
37.5	0.7	-2.8	-3.6	-2.7	-0.2	-0.7
34.0	0.5	-1.5	-2.1	-1.5	-0.0	-0.6
30.5	0.4	-0.9	-1.3	-0.9	0.0	-0.5
27.0	0.3	-0.6	-0.9	-0.5	0.1	-0.4
23.5	0.2	-0.3	-0.6	-0.4	0.1	-0.3
20.0	0.1	-0.3	-0.5	-0.3	0.1	-0.3
16.5	0.1	-0.2	-0.4	-0.3	0.1	-0.2

Note that the net heating rate is the sum of the solar heating, J_{SOL} in the first column, the solar infrared heating (not shown) and the total longwave, J_{LW}, in the third; the latter component is the sum of cooling rates due to the individual gases in the last three columns.

The net heating rate and the components are shown in table 4.3; there is net cooling at all heights, the solar heating reaching only $4.8Kday^{-1}$ at the stratopause compared to the $14.1Kday^{-1}$ cooling at the same level. The dominant contribution to the cooling arises from the CO_2, with lesser contributions from O_3 and H_2O, as discussed previously. The cooling varies in magnitude at different latitudes, but generally shows the same broad features. A point of interest is the level at which the 9.6μm band of ozone changes from giving heating to cooling; this is significant for the error analysis, since the magnitude of this heating rate is very small over a depth of a scale height and it is very sensitive to errors, in that a change of a fraction of $1Kday^{-1}$ shows as a large percentage error.

4.6.2. Sensitivity to the Tropospheric State.

Changes in the tropospheric ozone distribution lead to variations in the emitted radiation in the 9.6μ band, causing a change in the absorption of this radiation in the stratosphere. Perturbations of ±50% in the tropospheric ozone column led to changes of less than ∓10% in the heating in the low stratosphere, and generally less than 2% change at levels above 30km; at the changeover from heating to cooling the percentage error was much larger. The total longwave cooling was much less affected by these perturbations, changing by up to 15% below 20km but less than 1% above 25km; similarly, the net heating rate was quite insensitive to variations in tropospheric ozone distribution.

Tropospheric water vapour perturbations also produced a small response in the stratospheric cooling rates; changes of ±20% in the entire column caused changes in the water vapour cooling of less than ±5% above 20km and less than ±1% above 30km. Changes to the total cooling and net heating were also very small above 20km, a consequence of the minor role of water vapour in the radiation transfer above the low stratosphere.

Perturbations to the tropospheric temperature field cause changes in the radiance leaving the troposphere in the bands of all three trace gases considered here. Generally, the ozone and water vapour components responded in a manner similar to changing their tropospheric mixing ratios. The response in the 9.6μm band of ozone was slightly stronger for a ±3° change in the tropospheric temperature than for a change of ±50% in the mixing ratio. Changes in the CO_2 cooling for the same temperature perturbation were generally around 10% in the low stratosphere, but less than 1% above 25km and were negligible above 35km. The effect of these temperature perturbations on the total longwave cooling was also small above 25km (less than 5%) and negligible (less than 1%) above 35km, although the error could reach 20% at the lowest level (this could be exagerated due to the discontinuity in the temperature profile introduced by perturbing the tropospheric profile but not the stratospheric one, so it is likely to be a gross overestimate of the real error). Similar errors were evident in the net heating field as in the total longwave cooling.

The importance of clouds in the radiative balance of the low stratosphere

was emphasised by the work of Gille and Lyjak (1986), which showed that minima in the cloud distribution led to maxima in the ozone 9.6μm heating rate in the low tropical stratosphere; this gave a structure in the net heating field in the low stratosphere which was not evident in the calculations of Kiehl and Solomon (1986), who assumed clear skies in their calculations. In this thesis a zonal-mean cloud climatology was used, even though three-dimensional heating calculations were performed; in order to estimate the importance this in our calculations some experiments have been performed in which the cloud amount and height were perturbed. The results were very similar for each case, so only the former case will be discussed here.

The CO_2 cooling rates were very insensitive to the cloud distribution, but both the ozone and water vapour radiation transfer were affected. Assuming a cloud height of around 10km, increasing the coverage from 50% to 100% caused a decrease of around 60% in the 9.6μm ozone heating rate in the low stratosphere, which fell to around 25% in the mid stratosphere and less than 5% in the high stratosphere (with anomalously large percentage changes near the level of change from heating to cooling in the middle stratosphere). The response of the water vapour cooling was weaker, ranging from 40% near 20km to 1% at 35km. Because of the dominance of the CO_2 cooling with increasing altitude, these errors were reduced in the total cooling field, being around 50% at 16km, 10% at 30km and 1% at 45km; the net heating rate shows a similar sensitivity at high latitudes, where the cooling dominates, but in the tropics the stronger heating at lower levels causes the percentage change in the heating due to variations in the cloud field to be reduced, particularly above 25km.

In summary, the variations in the tropospheric state lead to small changes in the total cooling and net heating fields above 30km, but below this level the uncertainty in the troposphere can cause errors of perhaps 10% in each of these fields, and rather more below 20km. The errors in the components of the longwave cooling are rather stronger (particularly the ozone cooling below 30km, which is very sensitive to tropospheric temperature and cloud distributions).

4.6.3. Sensitivity to Errors in LIMS Data.

In order to place bounds on the errors due to inaccuracy in the LIMS data they were perturbed by their worst nominal accuracies, discussed in chapter 3. Entire columns have been perturbed; temperature was changed by ±2K, ozone by 15% at 10mbar to 40% at 100mbar and 0.1mbar (increasing linearly in log-pressure in these ranges) and water vapour by 25% throughout the stratosphere. These perturbations have been considered individually as well as considering the effect of a combined increase in ozone and decrease in temperature (and *vice versa*). Additionally, perturbations at single levels have been considered.

Changes of ±25% in the entire water vapour column resulted in changes of ±10-15% in the water vapour cooling throughout the middle atmosphere; there is an asymmetry to these changes, the decreased water vapour column leading to a stronger decrease in the cooling (typically 2%) than the corresponding increases. These changes show as a smaller perturbation in the total cooling and net heating fields, ranging from less than 10% in the low stratosphere to 1% at the stratopause (the response of the net heating being slightly stronger at high winter latitudes than in the tropics).

Ozone column perturbations affect both the 9.6μm cooling and the solar heating. Increasing (decreasing) the column results in increased (decreased) cooling above the middle stratosphere, with a change of ~30% (~25%) at the highest level; in the lower stratosphere the 9.6μm band heating is increased (reduced) by less than 10% (20%). The asymmetry of the response should again be noted. These errors are again much smaller in the total cooling (less than 5% at all levels). The response of the solar heating calculations to these ozone perturbations is quite large. For a mid-latitude profile an increase (decrease) in the ozone column causes an increase (decrease) in the heating of up to 30% below 30km, around 5% between 30-50km, but it can exceed 20% at high levels. When combined to give the net heating rate the changes partially cancel at high levels, where an increased ozone amount results in increases in the terrestrial emission and the solar heating; changes of up to 10% are found at high levels, but these can increase to around 30% near 16km. In the tropics, the balance in the net heating rate is somewhat different, with strong net heating at high levels which is sensitive to errors in the ozone data; the error in the net heating can reach 30%.

Perturbations of ±2K in the temperature column caused slight changes to the cooling profiles in middle latitudes; typically, for a positive perturbation the CO_2 cooling increased by 5%, ozone by 10% (stronger cooling at high levels and less heating at low levels) whilst water vapour cooling increased by 5–10% (stronger increase at low levels). The effect on the total cooling ranged from less than 5% at high levels to around 15% in the low stratosphere. A similar response was evident in the net heating in middle and high latitudes of the winter hemisphere, but again the dominance of solar heating in the tropics reduced the error in the net heating rate.

A coupled perturbation of temperature and ozone columns (increased temperature corresponding to decreased ozone) was also undertaken. The effect on the 9.6μm cooling was that it changed by less than in the case where only the ozone profile was perturbed, whilst at low levels, where this band heats, the error was increased. Accordingly, the total cooling was affected less at high levels but more at low levels than in the individual perturbations; the same is true of the net heating.

As may be anticipated, changes at an individual level produce a local response larger than that obtained by perturbing the entire column. The effect on the longwave cooling is to increase the exchange of heat with surrounding layers than from the deep perturbations. The response was very small more than a scale height from the perturbation.

CHAPTER 5

MONTHLY-MEAN DISTRIBUTIONS OF NET RADIATIVE HEATING RATE.

In this chapter the net radiative heating rates for the seven months of LIMS data are presented. Because the sum of the components (terrestrial and solar) does not provide a net field with a consistent global energy budget it is necessary to adjust it so that such balance is achieved. This was noted by Murgatroyd and Singleton (1961). Several recent studies of the zonal-mean net heating fields have used different methods of adjustment; these are discussed in the first section of this chapter. In section 2 our adjustment technique is presented. The zonal-mean heating fields are presented in section 3; they are compared to the previous publications. It is not straightforward to present three-dimensional fields of heating rates and convey all possible information; the presentation adopted here is by means of polar stereographic plots at isentropic levels which allow straightforward comparison with synoptic maps of potential vorticity. Two isentropic levels have been selected for presentation: η=7.5 and 6.7 (sections 4 and 5 respectively), which are close to 1mbar and 10mbar in the high and middle stratosphere.

5.1. Review of Previous Studies.

As in chapter 4 we commence our discussion of net radiative heating rates with a brief review of several recently-published cross-sections. Again, wherever possible, the fields for December will be examined. In each case the manner of adjustment of the heating field to obtain a globally consistent energy budget is discussed and the main features of the heating distribution reviewed: particular attention is given to the strength of the heating in the polar night, in the tropics and near the summer polar stratopause.

a. Gille and Lyjak (1986): LIMS data, December 1978. The net heating rates were adjusted so that the globally-averaged vertical flow of matter through each pressure surface vanished. An iterative method was used, based on the transformed Eulerian-mean thermodynamic equation:

$$\frac{\partial}{\partial t}[\theta] + \frac{[v]^R}{a}\frac{\partial}{\partial \phi}[\theta] + [w]^R\frac{\partial}{\partial z}[\theta] = [J_{NET}].$$

Assuming that the northward velocity vanishes everywhere a first

approximation to the vertical velocity was obtained from the zonally-averaged heating and potential temperature distributions. The constraint:

$$\int_{-\pi/2}^{\pi/2} \rho[w]^R \cos\phi \, d\phi = 0$$

was then applied, the heating rate being adjusted so that the global-average of the vertical velocity at each pressure level vanished. The iterative procedure then commenced: $[w]^R$ was used to calculate $[v]^R$ from the zonally-averaged continuity equation; this northward velocity field was used in the thermodynamic equation to re-estimate $[w]^R$; the adjustment process was then repeated and further iterations performed. Three iterations proved adequate to ensure convergence. This method of adjustment ensures that the northward heat transport is taken into account, so that the globally averaged net heating field itself need not vanish at each altitude. The global-mean heating and cooling profiles for 21 December 1978 are shown in figure 5.1a: a slight net heating is evident at high levels in the corrected heating field, indicating that the mean northward heat transport plays a role in the energy budget at such levels.

The net radiative heating rate for December (figure 5.2a) shows a maximum cooling rate exceeding $10Kday^{-1}$ above 0.4mbar (55km) in the polar night; this cooling rate decreases downwards, being $2Kday^{-1}$ at 3mbar (40km) and cooling only slightly at 100mbar. The zero-heating line is located at 30°N at 100mbar, sloping southwards and upwards to 10mbar, where it almost reaches the equator before sloping upwards and northwards, reaching 20°N at 0.5mbar. The net heating rate is positive over most of the tropics and the southern hemisphere, except for regions of weak cooling poleward of 60°S (which are in the region of extrapolated data) and in the low mesosphere. The strongest net heating rate exceeds $2Kday^{-1}$ at pressures lower than 2mbar at the equator and slightly south of this at higher levels.

b. Kiehl and Solomon (1986): LIMS data, December 1978. The unadjusted heating and cooling rates are not as well balanced at any level as those of Gille and Lyjak (compare figures 5.1a,b). Global energy balance was imposed by adjusting the global-mean cooling rate at each level so that it cancelled the solar heating; thus, no account is taken of the possible latitudinal variation of

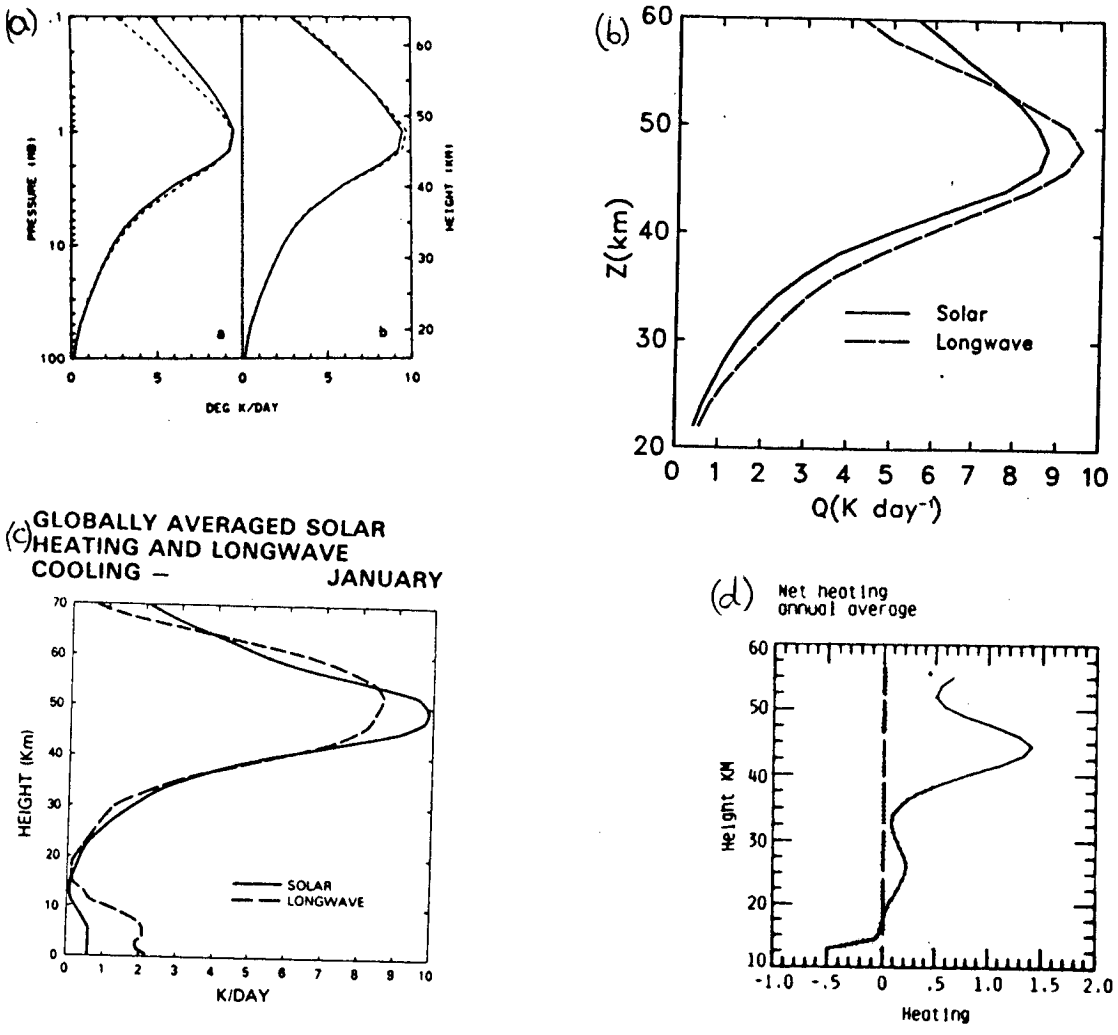


Figure 5.1. The globally averaged radiation balance as a function of height. (a) Solar heating (solid line) and terrestrial cooling (dashed line) before and after corrections to the net radiative heating rate, calculated by Gille and Lyjak (1986) using LIMS data from 21 December 1978. (b) Solar heating (solid line) and terrestrial cooling (dashed line) calculated by Kiehl and Solomon (1986) using January LIMS data. (c) as (b), calculated by Rosenfield *et al.* using climatological January data. (d) the annual average, calculated by Callis *et al.* using climatological data.

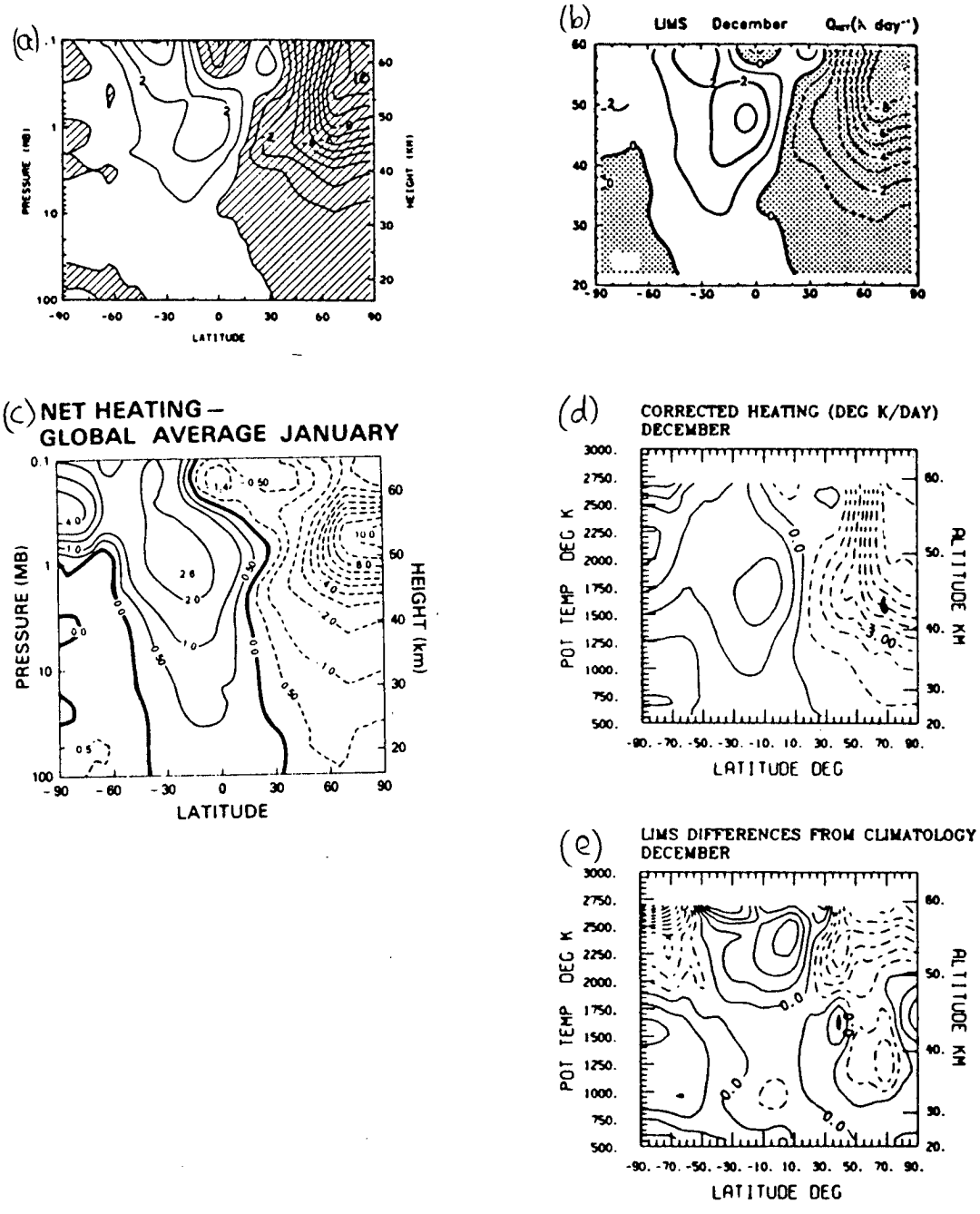


Figure 5.2. Corrected zonal-mean net radiative heating rates ($Kday^{-1}$) calculated by: (a) Gille and Lyjak, December LIMS data (contour interval $1Kday^{-1}$, cooling shaded); (b) Kiehl and Solomon (1986), December LIMS data (contour interval $1Kday^{-1}$, cooling shaded); (c) Rosenfield *et al.*, January climatology (contour interval $1Kday^{-1}$, with $\pm 0.5Kday^{-1}$ contours also shown, dashed contours indicate cooling); (d) Callis *et al.*, December climatological data (contour interval $1Kday^{-1}$, dashed contours indicate cooling); (e) Callis *et al.*, difference between LIMS and climatology for December (contour interval $0.4Kday^{-1}$, LIMS>climatology dashed).

static stability or of the meridional transport of heat. The adjusted net heating rates must therefore be regarded as less accurate than those of Gille and Lyjak. However, neglect of these quantities is consistent with the quasi-geostrophic approximation.

The net heating rate in December (figure 5.2b) shows cooling in excess of $8Kday^{-1}$ at 55km at the North pole; the $-2Kday^{-1}$ contour reaches the pole at 38km, slightly lower than in the study of Gille and Lyjak. The location of the zero heating line is close to that of Gille and Lyjak. The strongest heating at the equator exceeds $3Kday^{-1}$ near 50km altitude. In the summer hemisphere there is net heating above 44km and weak ($<1Kday^{-1}$) cooling at lower levels.

c. Rosenfield et al. (1987): Climatological data, January. Since December data are not presented by Rosenfield *et al.* we consider their January field, which can be expected to be less extreme than the previous month because of the different solar insolation. Again there are large differences between the unadjusted heating and cooling rates (figure 5.1c). The correction method was simply to subtract the global-mean of the unadjusted net heating rate from the heating rate at each latitude; no indication is given as to whether there was a latitudinal weighting in this procedure.

The adjusted heating field for January (figure 5.2c) shows maximum heating of $2.6Kday^{-1}$ near 1mbar at the equator and heating of $4Kday^{-1}$ near the South pole. There is net cooling almost everywhere below 1mbar and poleward of 40-50°S. In the northern hemisphere the zero-heating line is situated near 30°N between 100-1mbar, sloping equatorward at lower pressures. The strongest cooling is $10Kday^{-1}$ near 52km at the North pole. It should be noted that since a climatological data set was used the calculations cannot be expected to agree perfectly with the LIMS calculations of the previous two studies.

d. Callis et al. (1987): Climatological and LIMS data, December. This study presented heating rates with potential temperature as the vertical coordinate. In isentropic coordinates the natural adjustment to make to the net radiative heating field is to constrain the globally averaged vertical flux of matter across each isentropic surface to vanish (which is valid as long as we assume steady state); this constraint was applied. The correction to the net heating field was obtained from that of the net vertical mass flux field by working back through

the approximate relationship [J] = [W]$\Gamma_e c_p$, derived by Tung (1982).

The climatological net radiative heating rate (figure 5.2d) shows cooling of 9Kday^{-1} at the North pole at θ=2250K (η=7.7, z≈55km); the -2Kday^{-1} contour reaches the pole near θ=1000K (η=6.9, z≈33km). The zero-heating line is much further South than in the previously reviewed calculations. The strongest net heating rate of 2Kday^{-1} is slightly less than in the other studies, but these features are probably due to the use of a different data set and radiation transfer model. There is net cooling poleward of 50°S below the θ=1250K (η=7.1, z≈40km) isentropic level. The calculations using LIMS data (figure 5.2e) show 2Kday^{-1} stronger cooling than the climatology near the winter polar stratopause and the strongest heating is 1.2Kday^{-1} greater than in the climatological state; these corrections agree quite well with the calculations of Gille and Lyjak. In the southern hemisphere there is considerably less heating at higher levels when the LIMS data are used instead of the climatology and the cooling below 40km is weaker - they may give a slight net heating throughout the southern hemisphere stratosphere, again in good agreement with the results of Gille and Lyjak.

e. Overview. From this limited comparison it is evident that although the net heating distributions are qualitatively similar there are some quantitative differences between the four studies. The discrepancies between the results of Gille and Lyjak and those of Kiehl and Solomon, which were obtained from the same data but different radiation models, is a matter of some interest; if nothing else, it shows that there is still considerable progress to be made in our understanding of radiation transfer in the atmosphere. Consideration of the differences in the solar and terrestrial components of the net heating, discussed in chapter 4, suggests that there is incompatability between the radiation models used by these two studies and that it it this which causes the difference in the net heating fields rather than the different correction methods. The LIMS calculations of Callis *et al.* agree more closely with Gille and Lyjak than with Kiehl and Solomon.

5.2. Method of Adjustment: Global Balance.

In Callis *et al.* (1987) the globally averaged vertical mass-flux across each isentropic surface was constrained to be zero. This constraint is valid only if the mass above each isentropic level remains constant, which requires the

global-mean pressure at each level to be steady. Callis *et al.* did not attempt to prove this. Figure 5.3 shows the globally-averaged pressure of two isentropic levels, η=6.7 and 7.5, as a function of time between 1 November 1978 and 28 May 1979. It reveals that day-to-day changes in the pressure are very small, peaking during the periods of strong wave activity in the winter; even then the maximum changes in pressure at η=6.7 are 0.02mbarday^{-1}, which corresponds to a change of around 0.2% per day in the total mass above that surface. A similar maximum fractional mass change occurs at the higher level. Although there is a systematic change in the pressure above each isentropic level on a seasonal scale this too is very small, so the constraint that the global-mean mass-flow through each isentropic surface should vanish seems justified.

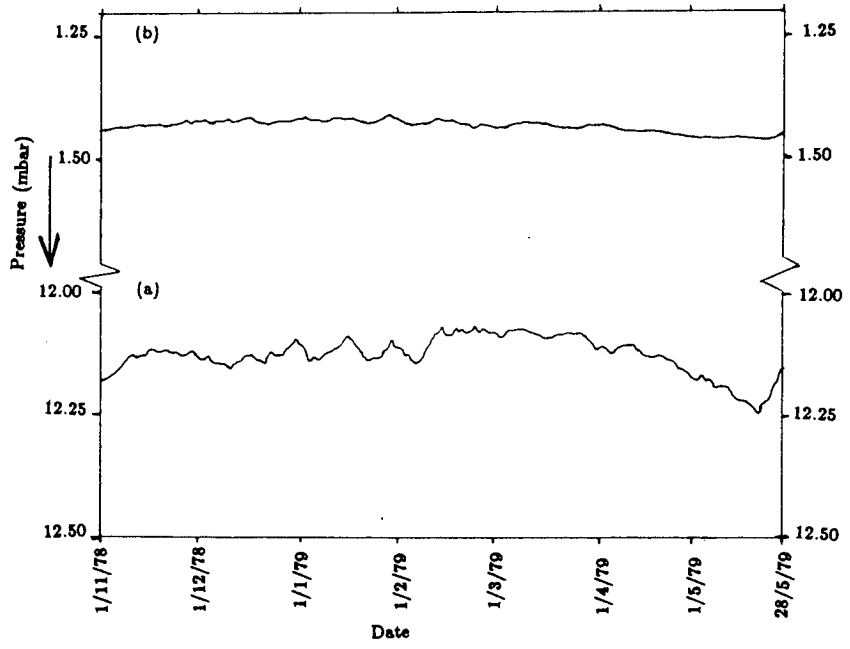


Figure 5.3. The global-mean pressure (mbar) of (a) the η=6.7 isentropic surface and (b) the η=7.5 isentropic surface from 1 November 1978 - 28 May 1979.

The approximations introduced by Tung (1982) and used by Callis *et al.* (1987) have not been applied in this study. Thus the vertical mass flux is defined by:

$$W_{UN} = \sigma_\eta J_{UN}/(c_p T)$$

(where the subscript UN denotes the unadjusted, J_{UN} being the sum of the solar and terrestrial components of the radiative heating rate). The zonal-mean, uncorrected, vertical mass-flux was thus determined from the three-dimensional fields of net heating rate, density and temperature. The zonal-mean heating rate, $[J_{UN}/c_p]$, for December 1978 is shown in figure 5.4. It is obvious that the cooling at all latitudes in the mid-stratosphere implies a globally-averaged decrease in the mean pressure of these isentropic levels, whilst a global-mean increase in the pressure of high isentropic levels is implied. The vertical mass-flux was thus adjusted so that it vanished at each isentropic level, in a similar manner to Callis *et al.* (1987). Because we are also interested in the three-dimensional distribution of net radiative heating rate and there is no evidence of a longitudinal bias in the original radiation transfer calculations, the adjustment to the mass-flux and radiative heating fields was assumed to be a function of latitude only.

The method of adjustment was to subtract a latitudinally-weighted residual of the global-mean mass-flux from the unadjusted flux at each grid point; it was derived by Dr. R.S. Harwood. If we denote latitude by the subscript j and longitude by i in the usual manner, the discrete form of the adjustment method was:

$$\{W_C\}_{i,j} = \{W_{UN}\}_{i,j} - \sum_j \{[W_{UN}]\}_j \cos\phi_j / \sum_j \cos\phi_j,$$

where the sum is over all latitudes and the subscript C denotes the corrected value. The corrected net radiative heating rate, J_C/c_p, was determined at each latitude using the definition of W and knowledge of the density and temperature fields. The zonal-mean heating distributions discussed in the following section and the three-dimensional fields of sections 4 and 5 were all corrected in such a manner. In the subsequent discussion the subscript C will be dropped from the corrected heating fields.

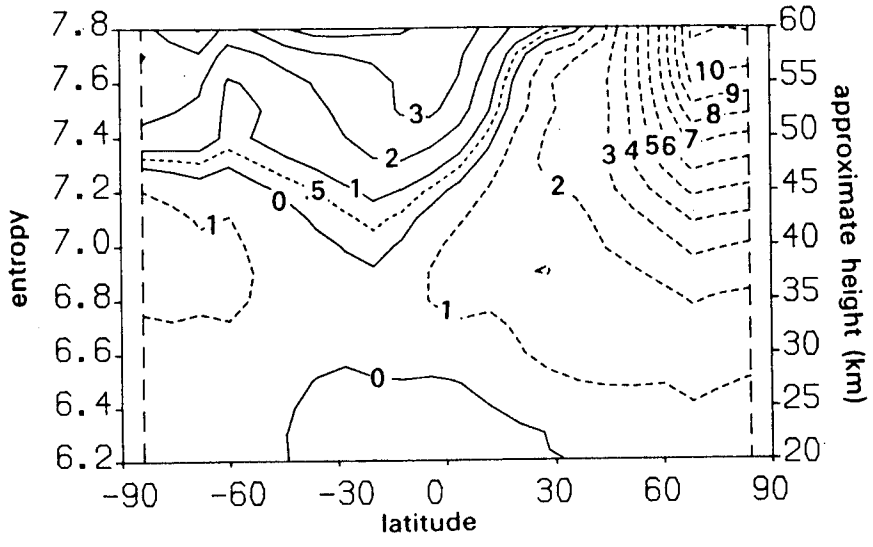


Figure 5.4. The unadjusted net radiative heating rate (J/c_p in $Kday^{-1}$) for December 1978. Regions of heating are denoted by solid contours and cooling by dashed contours, each with an interval of $1Kday^{-1}$; additionally, the $+0.5Kday^{-1}$ contour is shown (short dashes).

5.3. Monthly, Zonal-Mean Net Heating Distributions.

In this section the monthly, zonal-mean distributions of net radiative heating rate, $[J]/c_p$, are presented and discussed. These enable the latitude-height structure of the heating field to be understood; they can be used with the polar stereographic plots of the three-dimensional heating distribution to obtain a more complete understanding of the radiative budget of the stratosphere. Further, these zonal-mean cross-sections are the heating rates associated with the diabatic circulations presented in Chapter 6. Each month is discussed in turn.

a. November 1978 (Figure 5.5a). The strongest cooling in the northern hemisphere is $11Kday^{-1}$ near 76°N, η=7.8. The $-2Kday^{-1}$ contour is near 45°N at high levels but it slopes strongly downward and northward in the middle stratosphere, reaching the pole at η=6.8. The zero-heating contour is generally located in the northern hemisphere; below η=7.0 it is close to 30°N; above this level it slopes upward and equatorward, with slight cooling at the equator at

η=7.0; the line then slopes to 30°N at η=7.8. In the tropics the strongest net heating is slightly less than 2Kday^{-1} at 4°S, η=7.6; in the southern hemisphere there is a maximum of 4Kday^{-1} heating near 84°S, η=7.8. The lower southern hemisphere (η<7.3) cools poleward of 45°S, this cooling exceeding 1Kday^{-1} between η=6.6 and 7.3, poleward of 64°S.

b. December 1978 (Figure 5.5b). The net heating distribution for December is very similar to the previous month but more intense. There is still cooling at high latitudes and low levels in the southern hemisphere with heating above this (peaking at 3Kday^{-1}). The strongest heating in the tropics is 2.5Kday^{-1} (0-10°S, η=7.4-7.6). The zero-heating contour is in approximately the same location as in November, although the protrusion of heating into the northern hemisphere and cooling into the tropics near η=7.0 is slightly weaker this month. Cooling of 12Kday^{-1} occurs at the stratopause at the North pole. The -2Kday^{-1} line is closer to the equator at high levels and reaches the pole near η=7.0, rather higher than in November.

These heating distributions show the same qualitative features as the previous studies reviewed in section 1 of this chapter. The distribution of cooling calculated in the northern hemisphere is similar to that obtained by Gille and Lyjak, although our peak cooling rate is slightly stronger. In the tropics our results also agree fairly well with those of Gille and Lyjak, although they find slight cooling at high levels at the equator which we do not; agreement with Kiehl and Solomon is also quite good in these regions. In the region of extrapolated data (south of 64°S) our results at low levels are closer to those of Kiehl and Solomon than those of Gille and Lyjak. Perhaps the weakest point of this study is that the cooling near the South pole in the lower stratosphere is somewhat stronger than in the reviewed cross-sections, although our heating at higher levels near the South pole is weaker than all studies except that of Gille and Lyjak. Fels (1985) noted that since the summer polar stratopause is observed to be close to radiative equilibrium the net radiative heating rate should be small. Our results agree with this conjecture more closely than those of Kiehl and Solomon or Callis *et al*, whilst Rosenfield *et al.* find a net heating of 4Kday^{-1} at the south polar stratopause in January.

c. January 1979 (figure 5.5c). The strongest net heating rate at the summer stratopause reaches 3Kday^{-1} near 76°S, η=7.8, which is considerably weaker

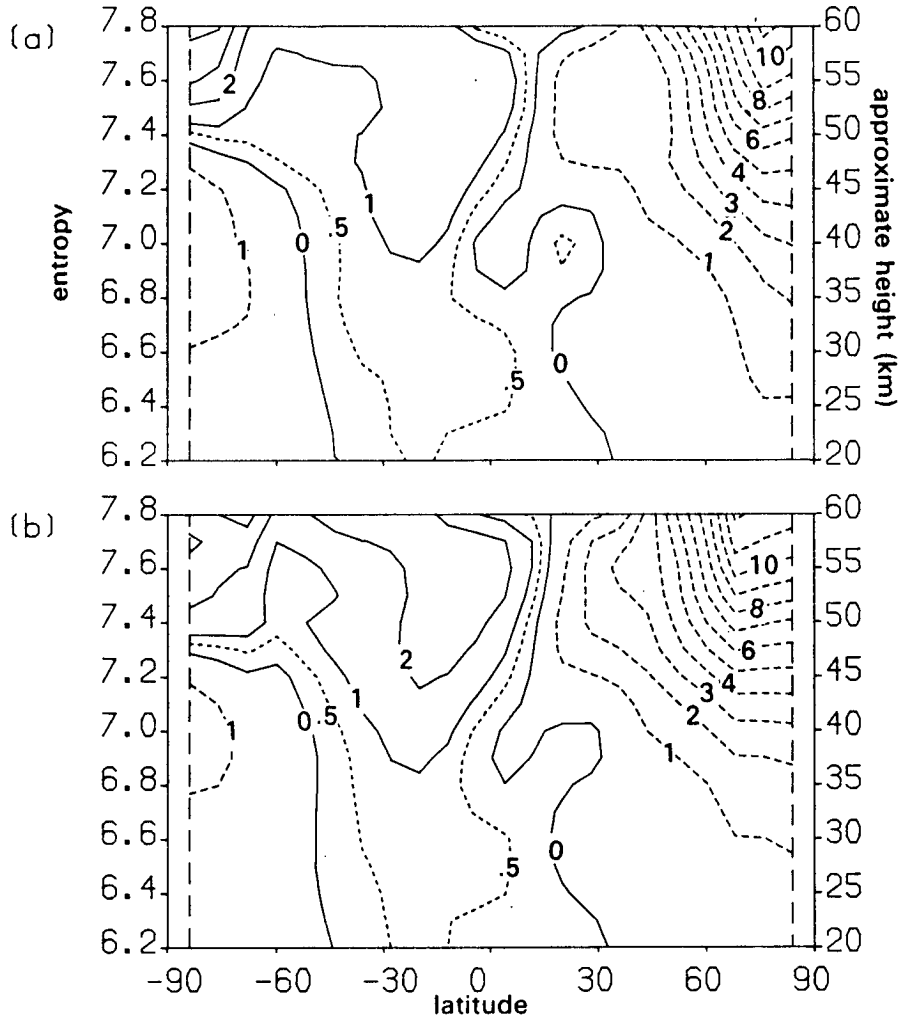


Figure 5.5. Net monthly, zonal-mean radiative heating rate (J/c_p in $Kday^{-1}$) for each month of the LIMS experiment. The contour interval is $1Kday^{-1}$; regions of heating have solid contours whilst cooling is indicated by dashed contours; additionally, the $+0.5Kday^{-1}$ contour is included (short dashes). The approximate height on the right-hand ordinate is defined as in Holton (1986) as $dz=\kappa^{-1}Hd\eta$, so that an interval of 0.1 in η is equivalent to 2.5 km; $\eta=6.2$ is assumed to be at 20km. This page shows (a) November 1978; and (b) December 1978. (Figure Continues.)

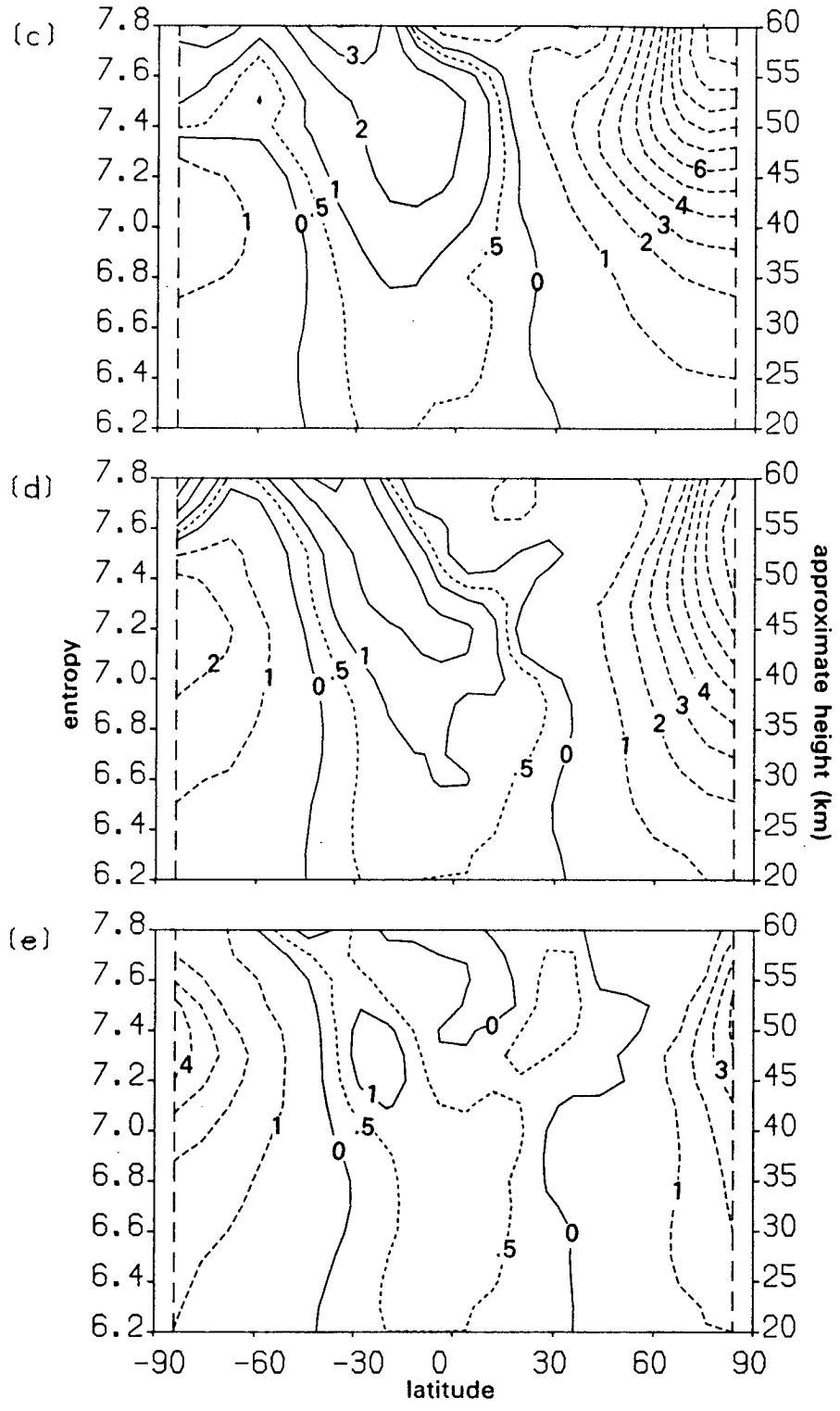


Figure 5.5 (continued). (c) January 1979; (d) February 1979; (e) March 1979. (Figure continues.)

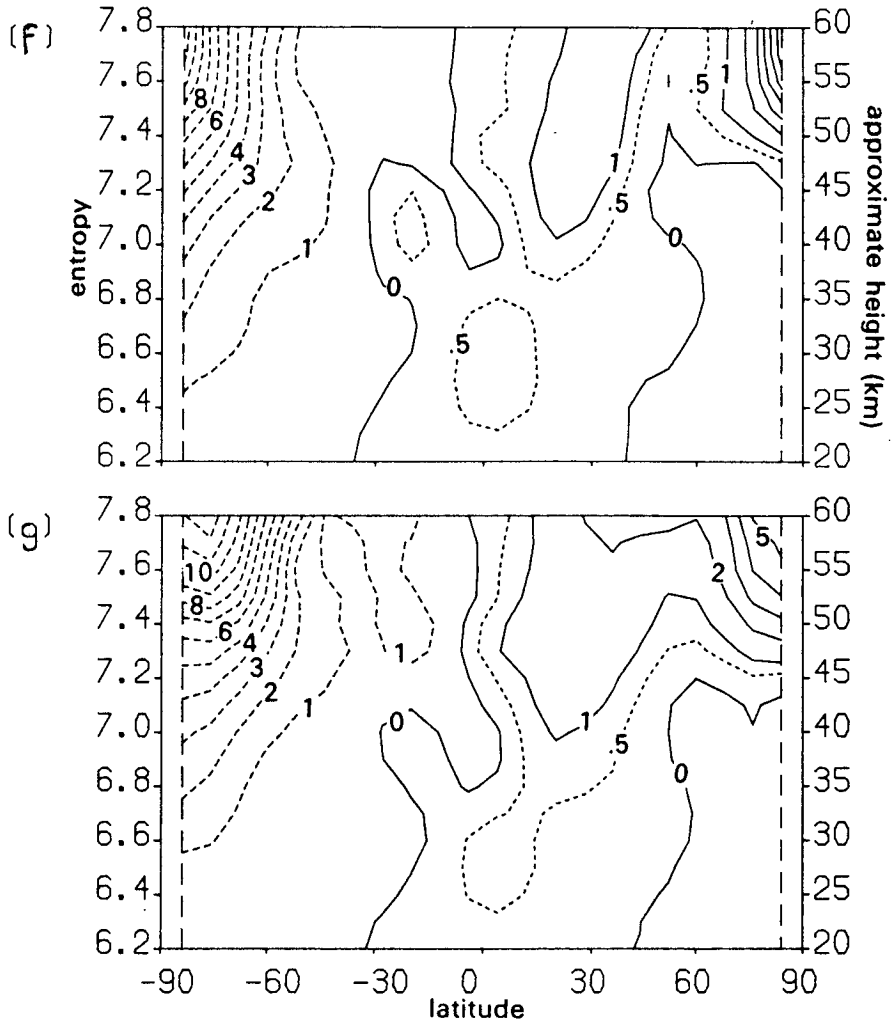


Figure 5.5 (continued). (f) April 1979; (g) May 1979.

than that obtained by Rosenfield *et al.* Near 30°S, η=7.8 the heating rate exceeds 3.5Kday^{-1}, decreasing downwards and equatorwards from this peak. At high equatorial levels there is more than 1Kday^{-1} cooling but below η=7.6 there is net heating; the zero-heating line slopes downwards and polewards from η=7.6, 15°N to η=6.7, 30°N. In the northern hemisphere the strongest cooling rate is 11Kday^{-1} above η=7.7 at the pole. The -2Kday^{-1} contour reaches the pole near η=6.8.

d. February 1979 (figure 5.5d). The illumination of the North pole as we come out of the polar night leads to a much weaker net cooling there than in the earlier winter months; the strongest cooling is now 8Kday^{-1} near η=7.6 and only 7Kday^{-1} at the highest level. The -2Kday^{-1} contour is much closer to the pole (near 60°N between η=6.7-6.8) but reaches the pole at η=6.6. The zero-contour is still near 30°N below η=7.0, but above this level it shows a more complex structure; there is cooling at the equator in much of the high stratosphere and low mesosphere. This region of net cooling, which propagated downwards through the mesosphere and high stratosphere between December and February was noted by Gille and Lyjak, who related it to the downward propagation of the semiannual oscillation, which was studied using LIMS data by Hitchman and Leovy (1986). There are two effects combining to cause stronger net cooling; the higher temperature results in stronger thermal cooling and, because at such levels a higher temperature implies a lower ozone mixing ratio (see Chapter 4), there is less solar heating (although this is complicated by the opacity argument, since photons are able to penetrate further into the atmosphere, so could be absorbed lower down, resulting in a stronger heating at slightly lower levels below the weaker heating above).

In the southern hemisphere tropics the heating rate exceeds 2.5Kday^{-1} above η=7.4. At southern high latitudes the strongest heating rate is 3.5Kday^{-1} at the pole at the highest levels. Poleward of 45°S there is net cooling at all levels below η=7.6; this cooling exceeds 2Kday^{-1} between η=7.0-7.4 south of 64°S, which is only slightly stronger than the cooling calculated by Gille and Lyjak.

e. March 1979 (figure 5.5e). In this equinoctal month the net radiative heating pattern shows a much more symmetric structure than in previous

months. The zero-heating line is positioned roughly symmetrically about the equator, although there is more structure in the northern than the southern hemisphere, presumably because the northern hemisphere has just been stirred up over the winter. The strongest net cooling is $4Kday^{-1}$ in each hemisphere, near η=7.4 at the poles. In the tropics there is a narrow band of cooling at high equatorial levels; this is flanked by a region of net heating in each hemisphere. The net heating rate exceeds $0.5Kday^{-1}$ between 20°S and 10-5°N in the lower and middle stratosphere, peaking at $1.5Kday^{-1}$ near 20°S, η=7.3.

f. April 1979 (figure 5.5f). The strongest heating is now in the northern hemisphere; a band in excess of $1.5Kday^{-1}$ is evident near 30°N, η=7.2. The heating rate approaches zero in middle latitudes, before increasing again poleward of 60°N; it exceeds $6Kday^{-1}$ near the North pole at high levels. There is cooling below η=7.3 poleward of 40-60°N. The zero-heating line in the southern hemisphere is close to 30°N below η=7.2, where there is a tongue of cooling penetrating from the southern hemisphere over the equator; at high levels the zero-heating line remains close to 4°S. The strongest cooling is $10Kday^{-1}$ near the polar stratopause.

g. May 1979 (figure 5.5g). The cooling at the winter polar stratopause reaches $12Kday^{-1}$, although this is in the region of extrapolated data, so must not be regarded as too accurate. The $-2Kday^{-1}$ contour is at 45°S η=7.8 and reaches the pole at η=6.8; it is very similar in location and orientation to its northern hemisphere counterpart in November 1978 (figure 5.5a). Likewise, the zero-heating contour shows an interhemispheric, seasonal symmetry. The high-level heating at the North pole in May is $5Kday^{-1}$, stronger than at the South pole in November. The zero-heating level in the summer hemisphere is lower (η=7.2) in May than in November (η=7.4) and the cooling in the summer polar lower stratosphere is weaker in May than in November. Such differences at high latitudes should be interpreted with caution because of the extrapolation of the data in this region of the southern hemisphere.

5.4. Net Radiative Heating Distribution at η=7.5

Attention is now devoted to a presentation of the three dimensional fields of net radiative heating rate for each month of LIMS data. Two isentropic levels have been selected; discussion of the heating distributions are given

along with a description of the vortex structure, defined in terms of the monthly-mean potential vorticity distribution.

In this section we consider the η=7.5 isentropic level, the global-mean pressure of which remained close to 1.2mbar throughout the LIMS observation period (figure 5.3b). Polar stereographic plots of potential vorticity (20-80°N) and net radiative heating rate (4-84°N) are presented.

a. November 1978, figure 5.6. Poleward of 60°N both the potential vorticity and the net heating rate showed only weak nonzonality; this is because of the tendency of the planetary waves to propagate upwards and equatorwards in the early winter, leaving the polar cap relatively undisturbed. Between 20-60°N there was some zonal-asymmetry, particularly in the heating distribution; the zonal-mean net heating rate at 40°N was $-1.5 Kday^{-1}$ but the three-dimensional field showed a region of slight net heating around 180°E and cooling in excess of $3 Kday^{-1}$ at 300-330°E.

b. December 1978, figure 5.7. Analysis of the net heating distribution is simplified if the solar and terrestrial components are considered first (figure 5.8). Since there is little solar insolation in the winter polar cap the solar heating was close to zero (figure 5.8b); the heating rate increased with decreasing latitude, any non-zonalities being well correlated with the ozone distribution at this level (figures 5.8a,b); of particular significance is the longitudinal maximum in heating associated with the ozone maximum between 90-150°E, 20-60°N. Because the longitudinal maximum of ozone corresponds to a minimum in temperature at this level (as discussed in chapter 4; figures 5.8a and c), the weakest terrestrial cooling coincided with the strongest solar heating and the non-zonality was enhanced in the net heating field. Thus the net heating distribution (figure 5.7b) showed a region of very slight heating near 150°E, 40°N, compared to the $3 Kday^{-1}$ net cooling at the opposite side of the pole. The strongest net cooling in the cap exceeded $9 Kday^{-1}$ and was situated to the east of the potential vorticity maximum, which itself was to the east of the temperature minimum. Equatorwards of 16-20°N there was net radiative heating, which approached $2 Kday^{-1}$ at the equator.

c. January 1979, figure 5.9. The polar vortex was more strongly displaced than in the previous month, its centre being near 330°E, 75°N, with a minimum of potential vorticity near 150°E, 55°N. The wavenumber-1 displacement was

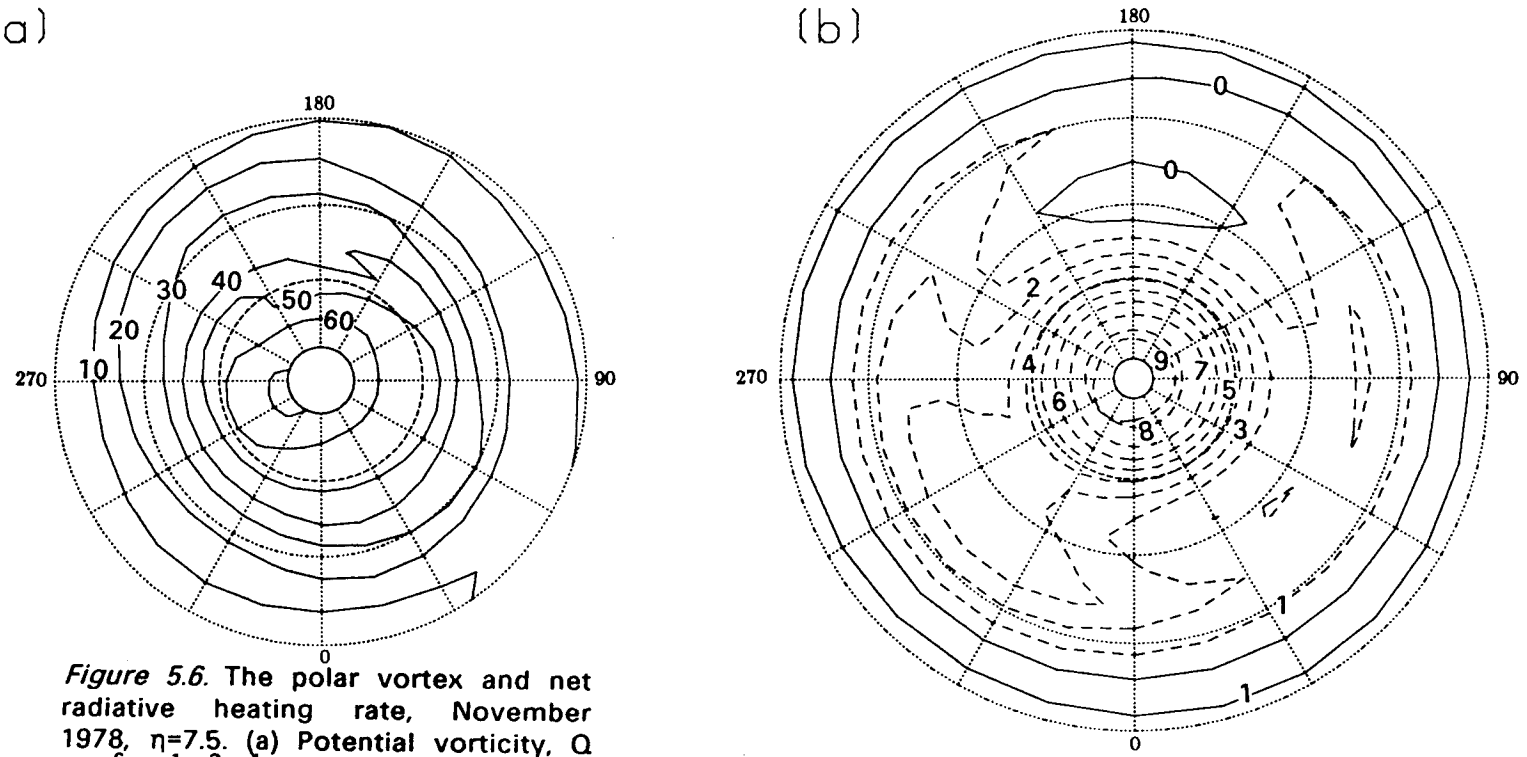


Figure 5.6. The polar vortex and net radiative heating rate, November 1978, η=7.5. (a) Potential vorticity, Q ($10^{-6}kg^{-1}m^2s^{-1}$) calculated from LIMS data (contour interval $10\times10^{-6}kg^{-1}m^2s^{-1}$) between 20°N and 80°N; (b) net radiative heating rate, J/c_p ($Kday^{-1}$, contour interval $1Kday^{-1}$, solid lines indicate heating, dashed lines indicate cooling) between 4°N and 84°N. (These are polar stereographic projections.)

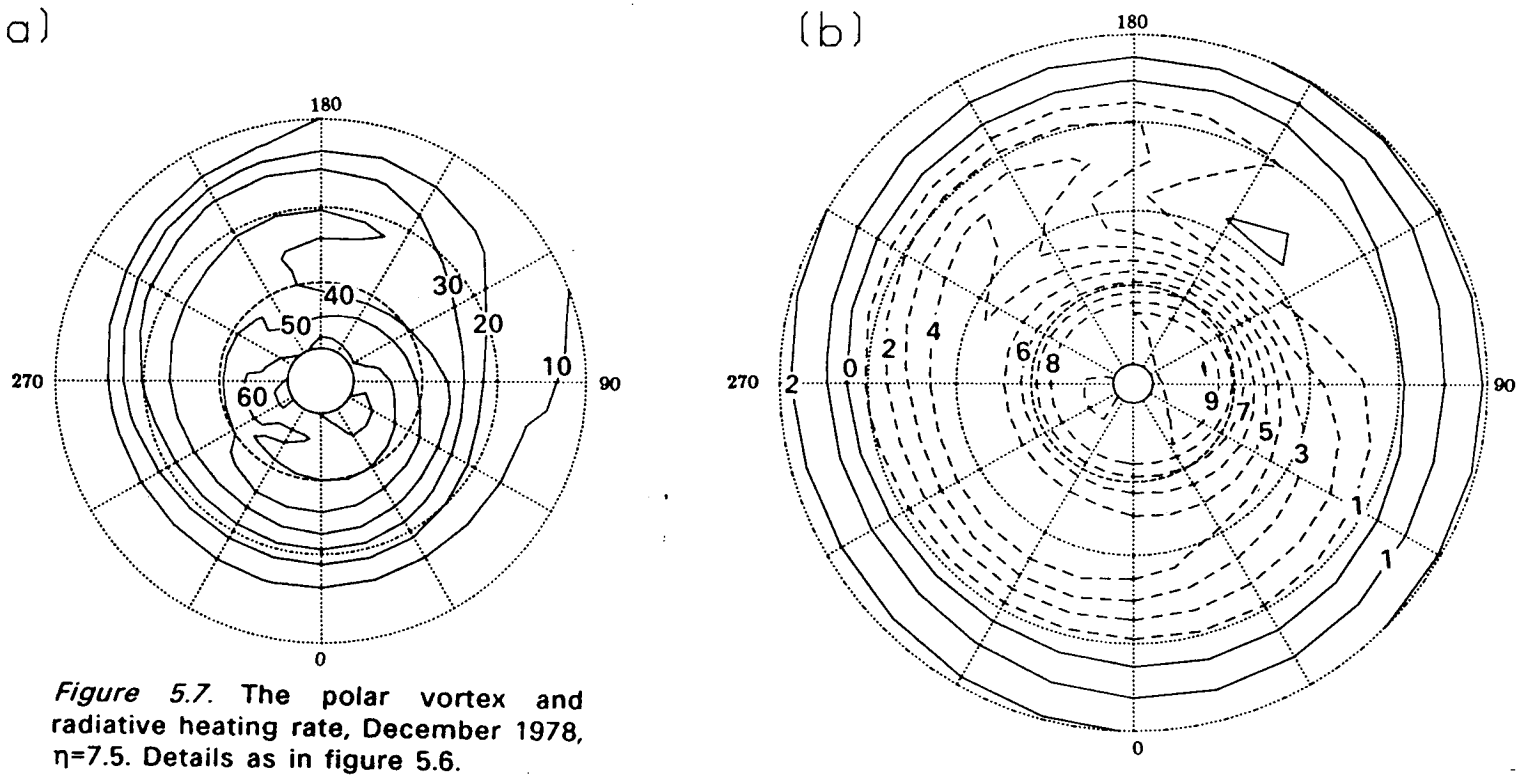


Figure 5.7. The polar vortex and radiative heating rate, December 1978, η=7.5. Details as in figure 5.6.

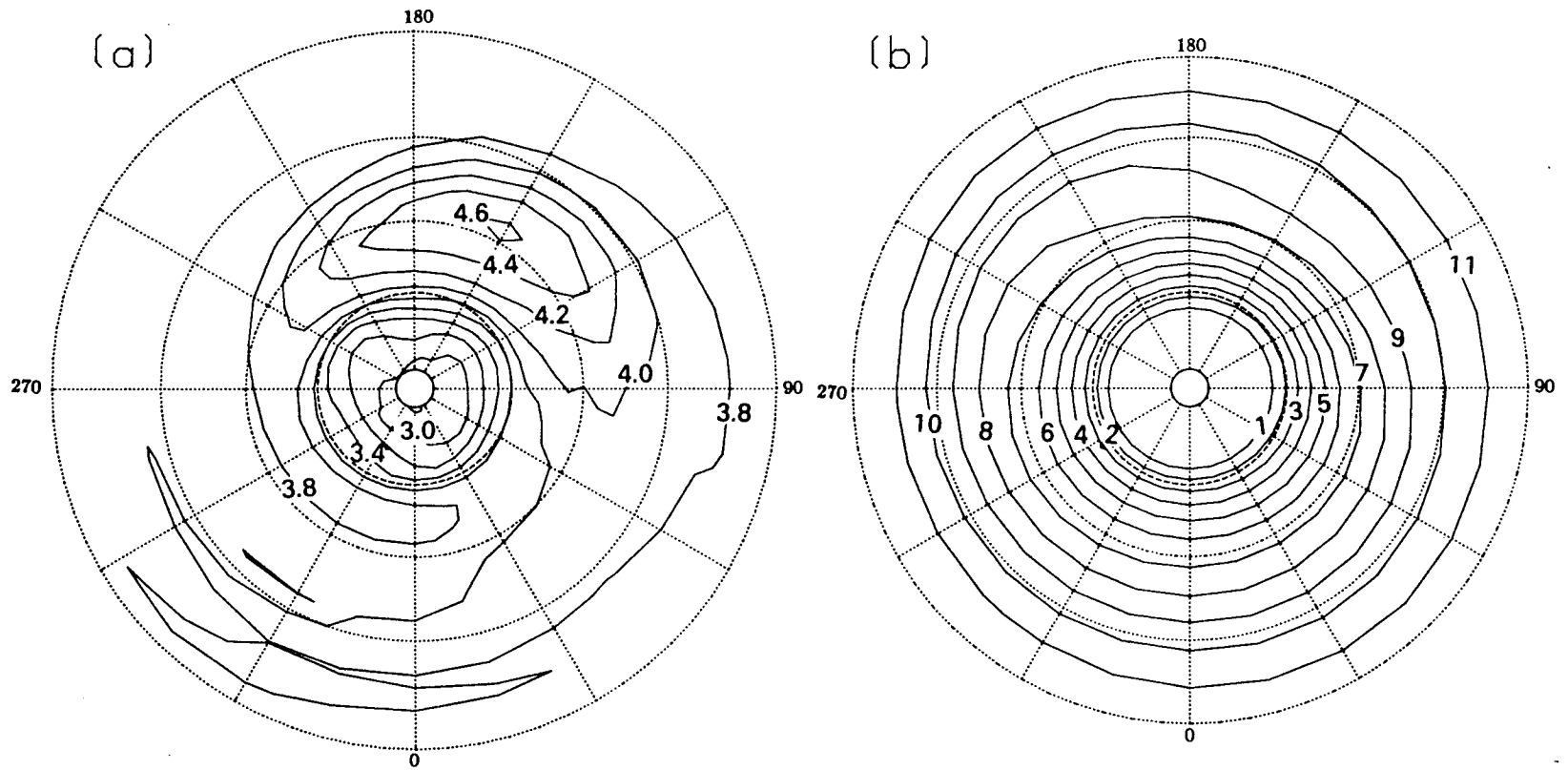
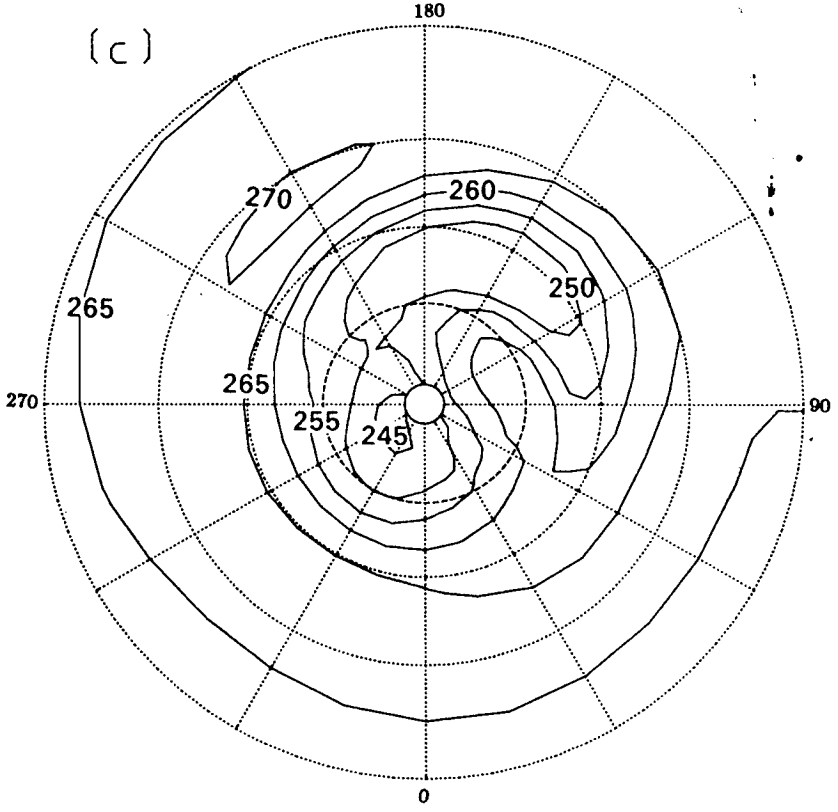


Figure 5.8. Polar stereographic plots at η=7.5, December 1978. (a) LIMS ozone (ppmv, contour interval 0.2ppmv) between the equator and 84°N; (b) solar heating rate (J/c_p, $Kday^{-1}$, contour interval $1Kday^{-1}$) between 4°N and 84°N; (figure continues).



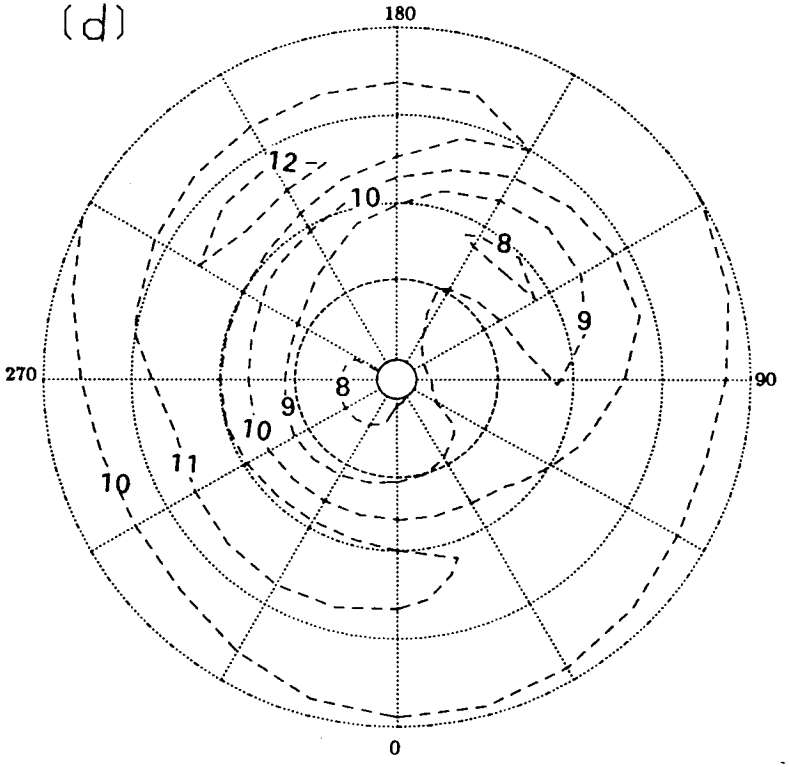


Figure 5.8 (continued). (c) LIMS temperature distribution (K, contour interval 5K) between 0° and 84°N; (d) terrestrial cooling rate ($Kday^{-1}$, contour interval $1Kday^{-1}$) between 4°N and 84°N.

also evident in the net heating pattern, which was more intense than in December; the monthly mean temperature distribution had a cold centre displaced from the pole to 260°E, 75°N with a warm pool on the opposite side of the pole over Siberia (on the eastern flank of the vortex). Strong cooling in this warm air with weaker cooling over Canada combined with the almost zonal solar heating pattern to give a wavenumber-1 pattern in the net heating rate; the strongest cooling, 11Kday^{-1}, being over Siberia at 75°N. Again, there was heating equatorward of 15-20°N.

d. February 1979, figure 5.10. A strong wavenumber-1 pattern persisted, with the vortex slightly to the East of its January position. The peak cooling rate of 9Kday^{-1} was centred over Scandinavia. Of particular interest this month is that between 20-60°N there was slight net heating over the Pacific and North America but strong cooling over Europe; interpreted in terms of vertical velocity, this implies weak upward motion through a large area of the isentropic surface but strong downward motion in other regions. Note the net cooling at all tropical longitudes, which was also evident on the zonal-mean cross section, discussed in the previous section.

e. March 1979, figure 5.11. Progressing into spring the polar vortex became weaker and more zonally symmetric but still showed a slight wavenumber-1 structure. The increased solar insolation at the pole caused the net radiative cooling to decrease to 4Kday^{-1}. The net radiative heating rate exceeded 1Kday^{-1} between 170-270°E, near 40°N and was generally weak and positive between 20-40°N with cooling in most of the tropics.

f. April 1979, figure 5.12. There was net heating over most of the northern hemisphere with a small region of net cooling between 280-90°E, 40-60°N. The net heating rate was strongest at the North pole, where it reached 4Kday^{-1}, and in the tropics (>2Kday^{-1}). The weak zonal-asymmetry was consistent with the relatively undisturbed polar vortex.

g. May 1979, figure 5.13. The polar vortex showed only weak non-zonalities. The net heating rate exceeded 1Kday^{-1} over almost all of the northern hemisphere. It was stronger than 2Kday^{-1} in two regions in middle latitudes and reached 4Kday^{-1} at the North pole.

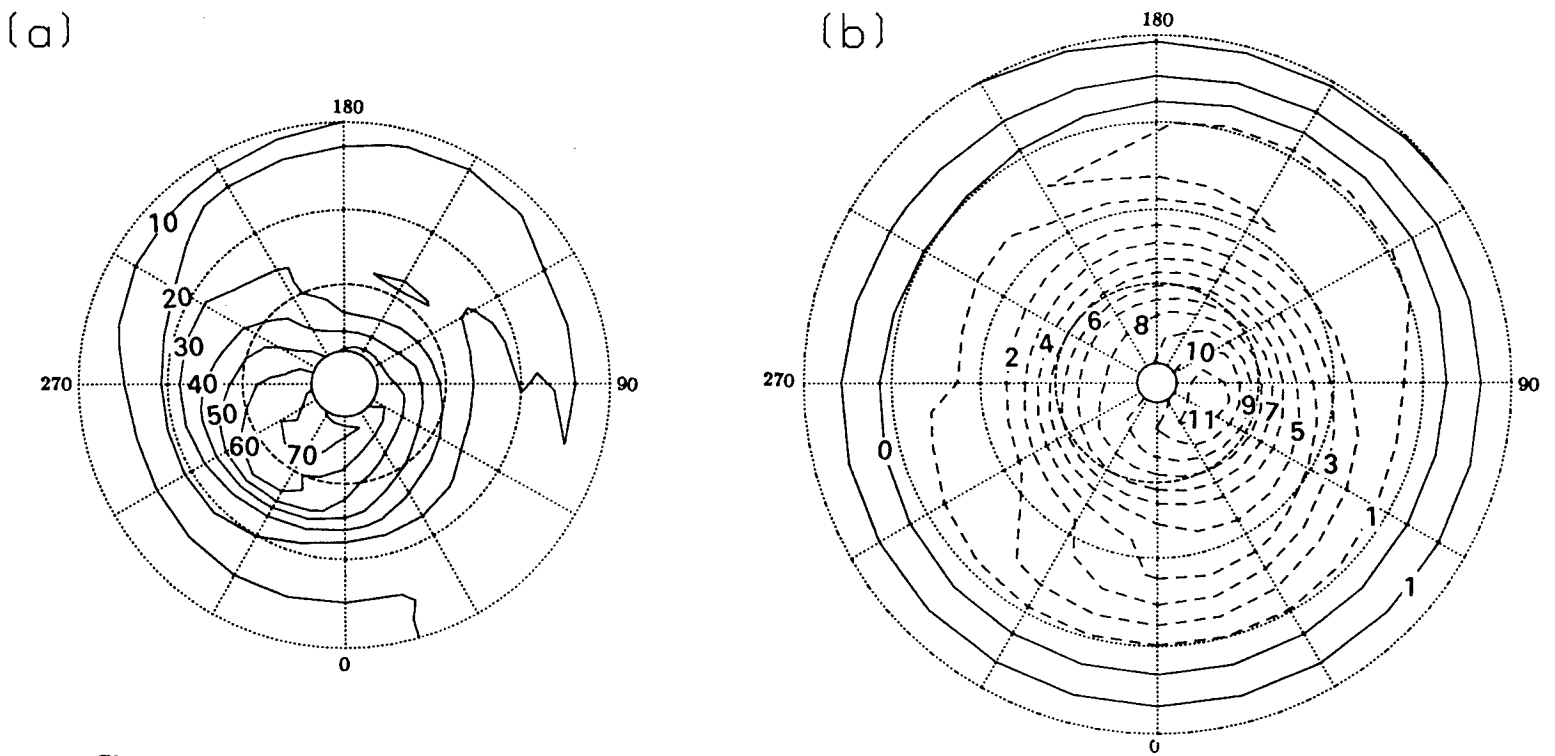


Figure 5.9. The polar vortex and radiative heating rate, January 1979, η=7.5. Details as in figure 5.6.

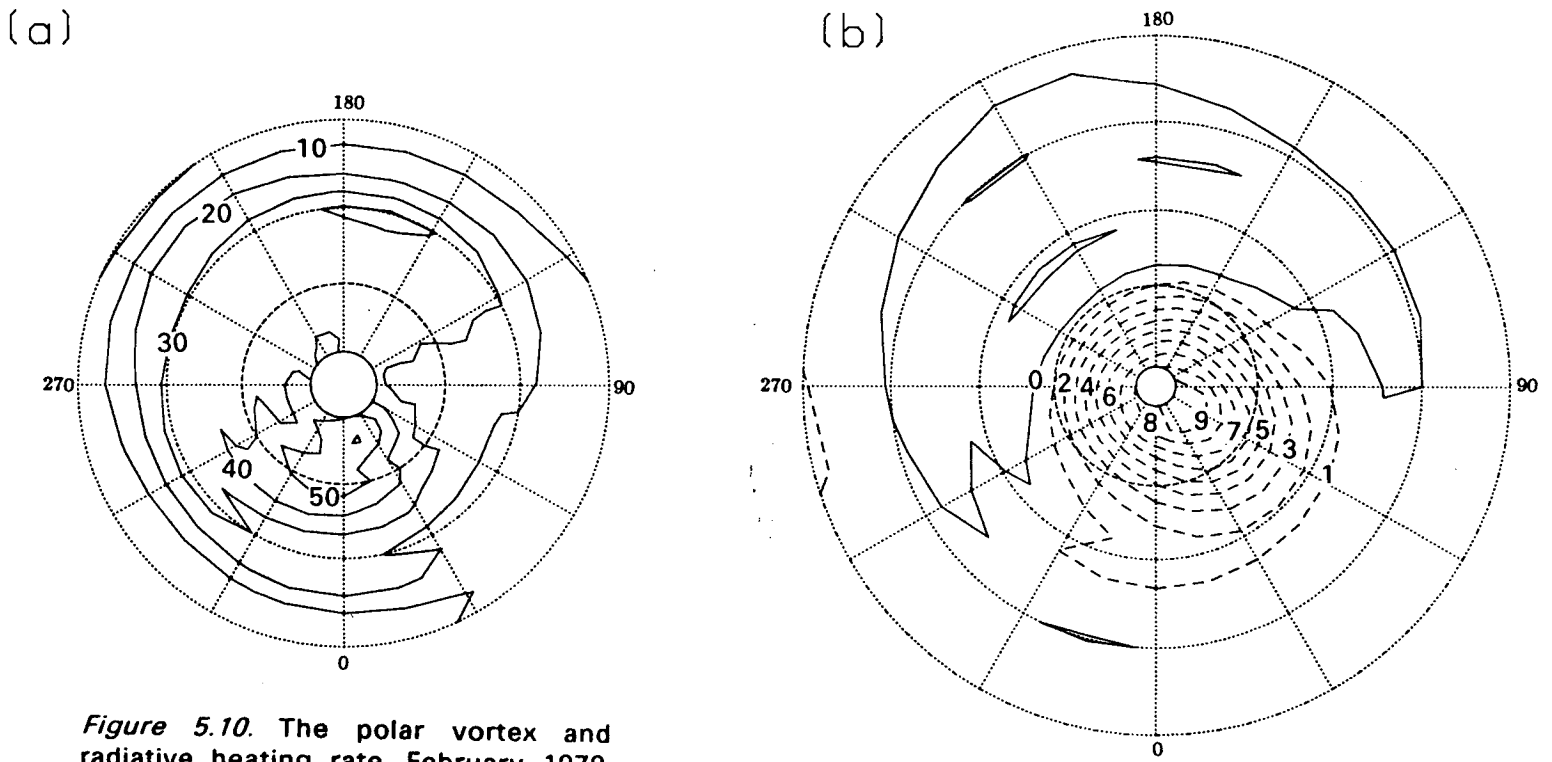


Figure 5.10. The polar vortex and radiative heating rate, February 1979, η=7.5. Details as in figure 5.6.

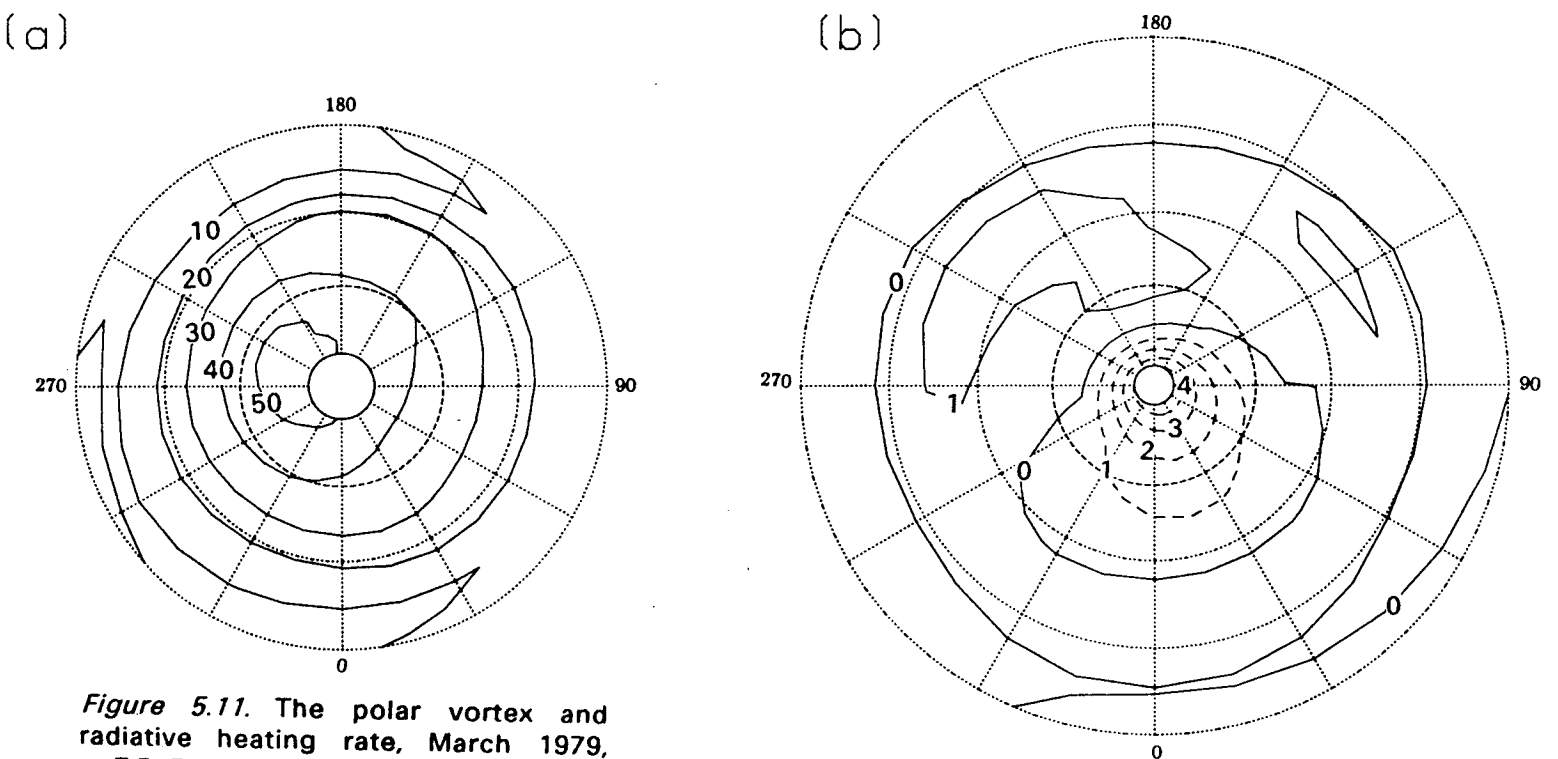


Figure 5.11. The polar vortex and radiative heating rate, March 1979, η=7.5. Details as in figure 5.6.

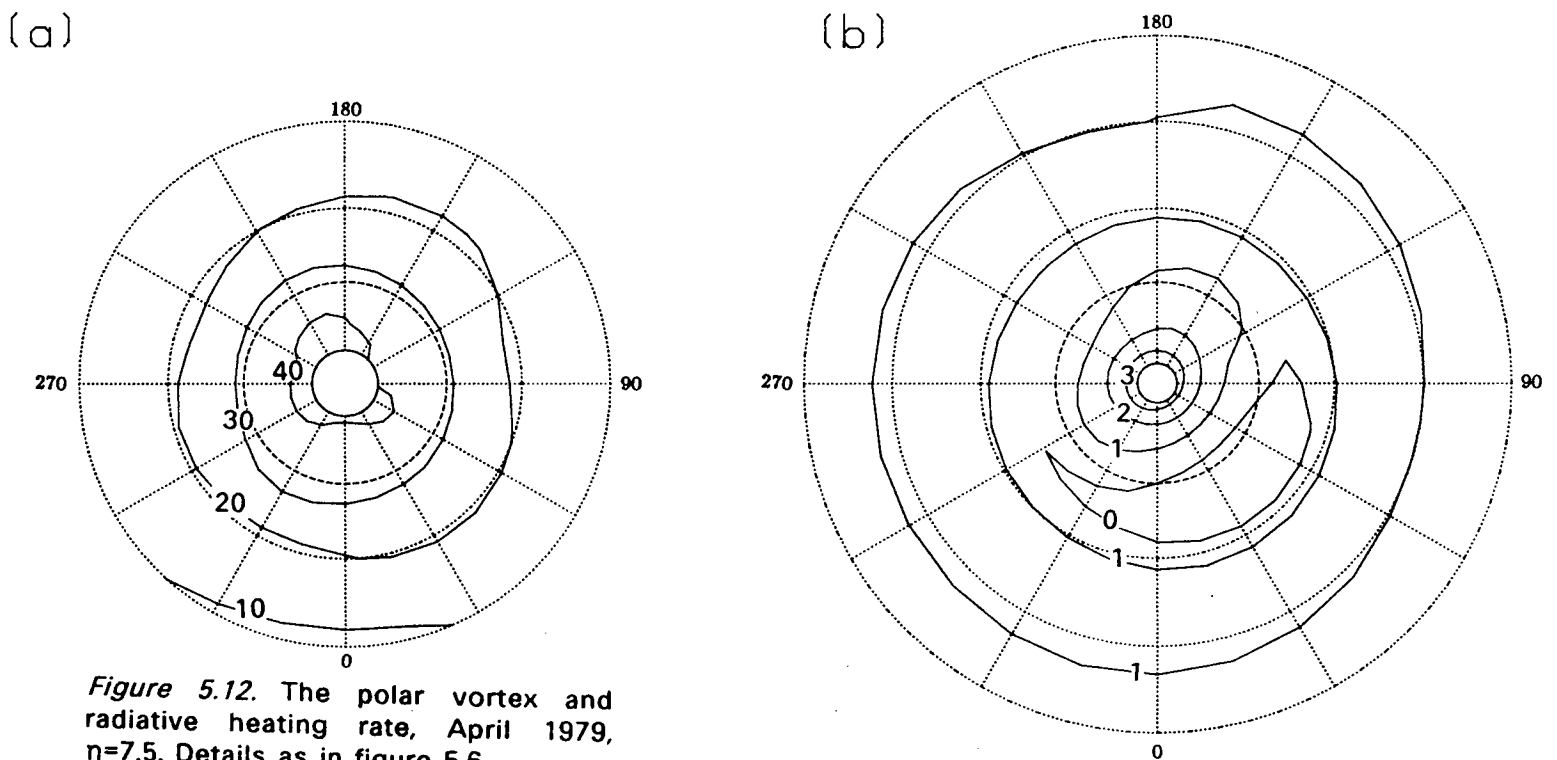


Figure 5.12. The polar vortex and radiative heating rate, April 1979, η=7.5. Details as in figure 5.6.

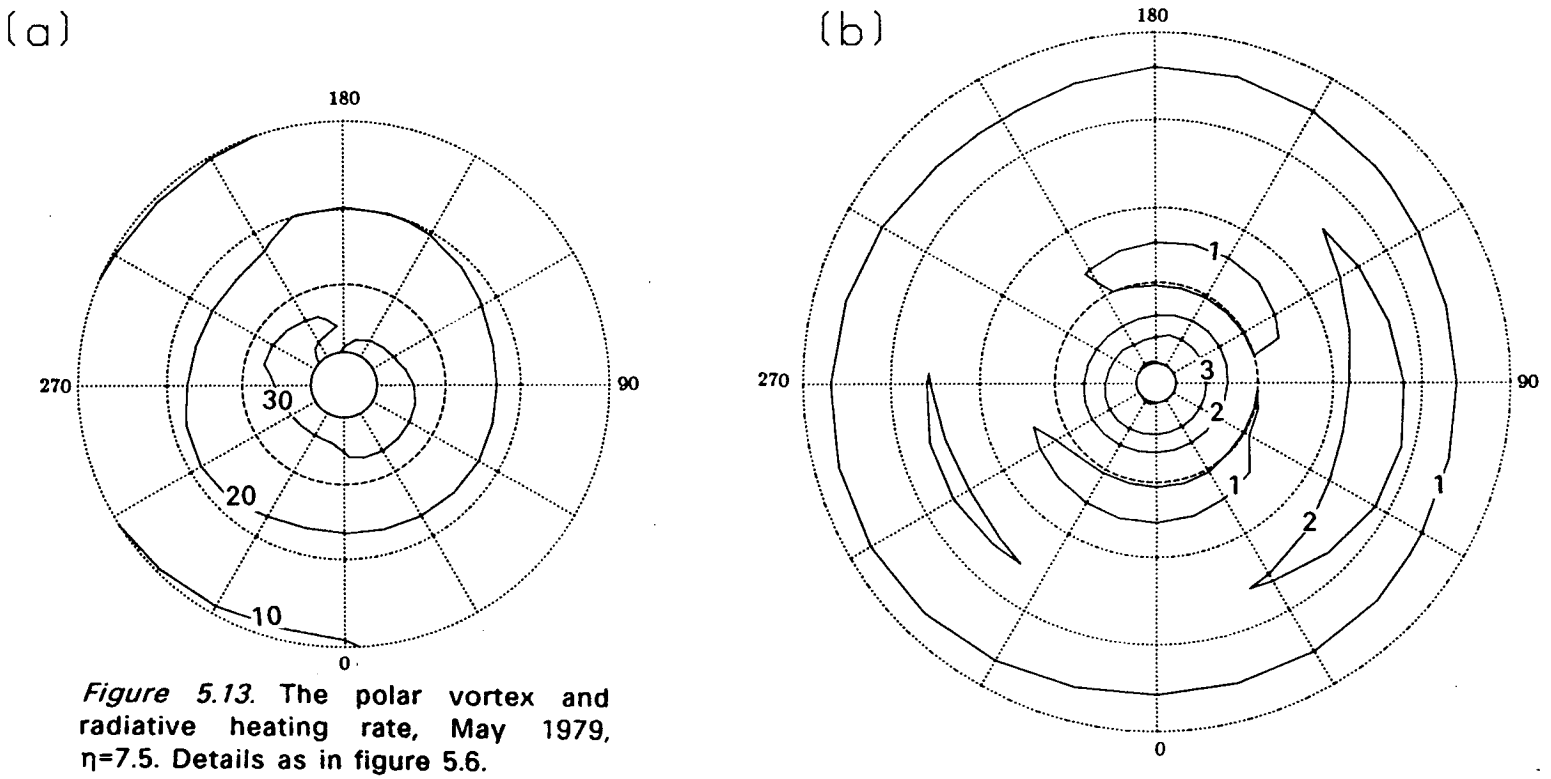


Figure 5.13. The polar vortex and radiative heating rate, May 1979, η=7.5. Details as in figure 5.6.

5.5. Net Radiative Heating at η=6.7

Monthly-mean distributions of the net radiative heating at η=6.7 are presented in this section. Again we present fields for the northern hemisphere between December 1978 and May 1979. The potential temperature of this isentropic level is 812K; thus it is close to the 850K level selected by McIntyre and Palmer (1984) and others for studies of the nonlinear dynamics of the stratosphere. The global-mean pressure of this surface remains close to 12mbar throughout the winter of 1978-79 (figure 5.3a). Because of the strongly nonlinear, rapidly changing dynamical activity in the winter middle stratosphere, standard deviations of the daily fields of potential vorticity and radiative heating rate from the monthly-mean distributions are considered.

At this level both the heating and cooling rates are weaker than near the stratopause. Generally, the net heating rate showed cooling in middle and high latitudes and heating in the tropics (see the zonal-mean cross sections of figure 5.5).

a. November 1978, figure 5.14. The potential vorticity distribution shows a wavenumber-1 distortion, the vortex centre being near 270°E, 80°N. There was an associated weak zonal-asymmetry in the radiative heating distribution. There was longitudinal structure in the heating field between 30-60°N, where the potential vorticity contours were also buckled.

b. December 1978, figure 5.15. A more strongly displaced wavenumber-1 vortex was evident this month, the centre being at 75°E, 75°N. The zonal asymmetry in the net radiative heating field was also enhanced, the strongest net cooling (1.8Kday^{-1}) being situated at 135°E, 50°N to the East of the polar vortex. The dominant pattern is wavenumber-1, but there were contributions from higher wavenumbers, particularly near 40°N. The net heating in the tropics was weak, reaching 0.4Kday^{-1} in some regions, and also showed slight non-zonality.

c. January 1979, figure 5.16. The general structure of potential vorticity and heating rate was similar to that of December 1978, although the extrema of both fields were intensified. The strongest cooling, exceeding 2.4Kday^{-1}, near 120°E, 60°N was again to the East of the vortex where the warm tropical air was advected by the cyclonic flow. A tongue of high potential vorticity

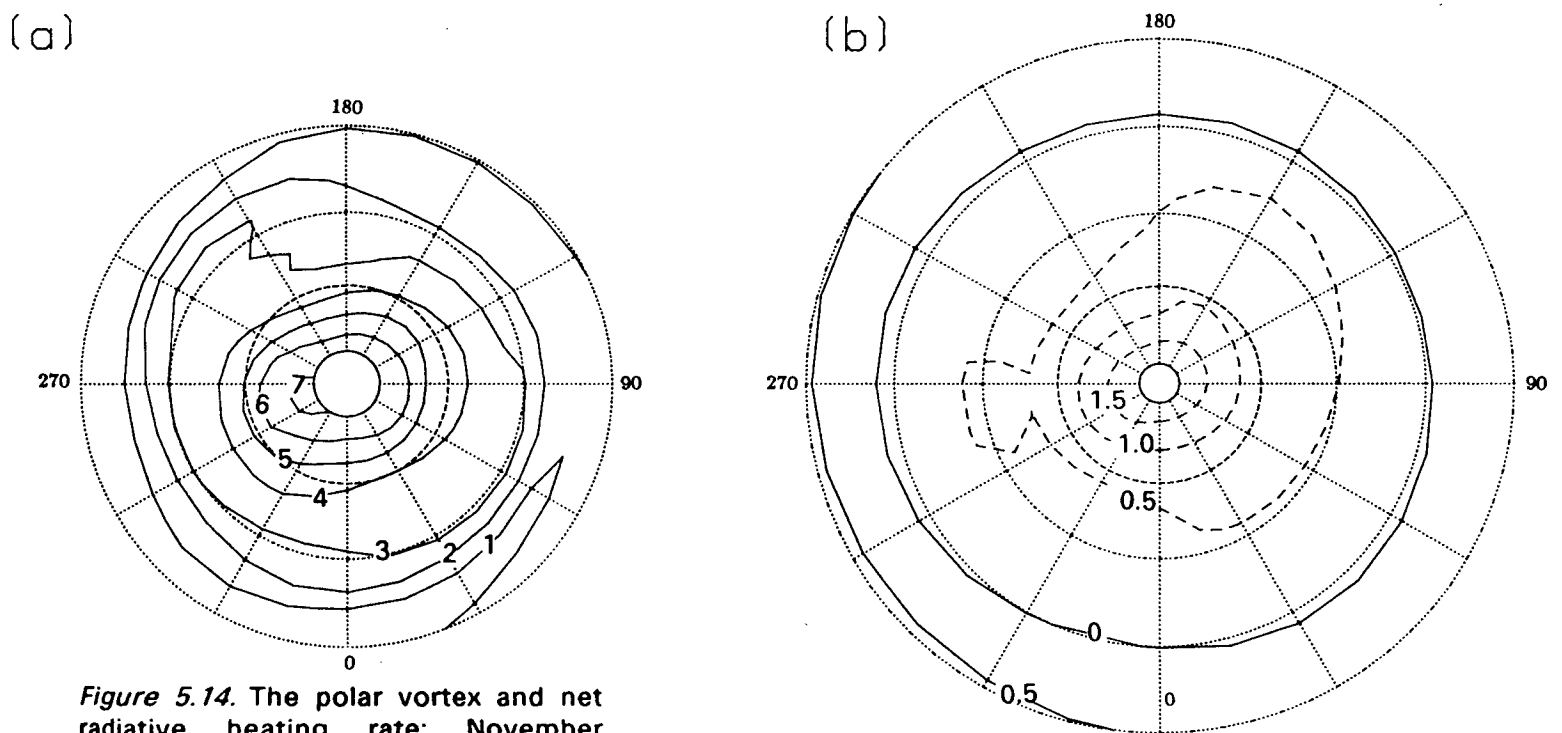


Figure 5.14. The polar vortex and net radiative heating rate; November 1978, η=6.7. (a) potential vorticity (10^{-6}kg^{-1}m^{2}s^{-1}, contour interval 1×10^{-6} kg^{-1}m^{2}s^{-1}) between 20°N and 80°N; (b) net radiative heating rate (Kday^{-1}, contour interval 0.5Kday^{-1}) between 4°N and 84°N.

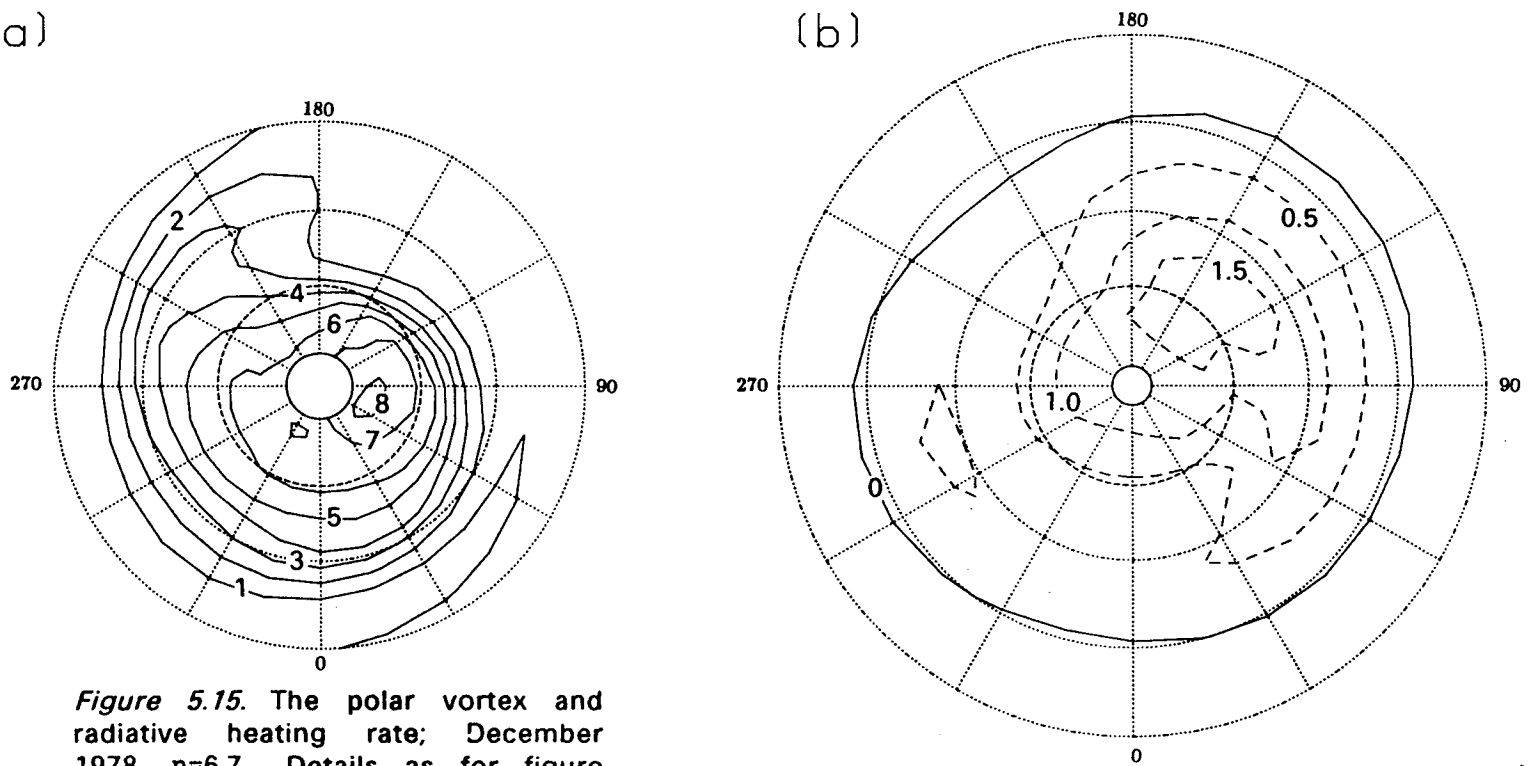


Figure 5.15. The polar vortex and radiative heating rate; December 1978, η=6.7. Details as for figure 5.14.

extended from the eastern flank of the vortex, coincident with a secondary longitudinal minimum of net radiative heating near 300°E; this gave a wavenumber-2 structure to the heating field near 40°N.

The extent of the non-zonal, transient wave activity in the northern winter stratosphere in the winter of 1978-79 has been the topic of much discussion. A strong wavenumber-1 (with a minimum in wavenumber-2) dominated the temperature fields in January (Labitzke, 1981); this wave grew to extremely large amplitude between 16 and 26 January, when a minor warming occurred. This amplification process was interpreted as an interaction between a forced stationary wave and a free travelling wave by Madden and Labitzke (1981). The nonlinear state of the atmosphere at this time can be recognised by examination of the daily distributions of potential vorticity (Dunkerton and Delisi, 1986), a time series of which shows the vortex moving off the North pole in the second half of January and distorting as the warming occurred, large patches of 'surf' being torn from the main vortex (McIntyre and Palmer, 1983, 1984).

The tongue of high potential vorticity air near 150°E, 25°N on the January monthly-mean plot (figure 5.16b) was evident on the daily charts at the beginning of the month, when a large, wave-3 vortex showed signs of erosion in this area; also, tongues of high potential vorticity air passed this region as they were advected from the western flank of the distorted vortex at the end of the month. The standard deviation (figure 5.16c) shows a region of some variability slightly to the north of this feature where the vortex debris also left its signature. The region of strong standard deviation at 60°E, 55°N is coincident with the location of the vortex centre when it was displaced from the pole between 19-25 January, whilst that at 260°E, 70°N arises because this region was enveloped in the vortex when it was centred on the pole at the beginning of the month. Variations near 40°N show where wave-breaking occurred. The monthly-mean potential vorticity distribution (figure 5.16b) is thus an average of two distinct dynamical situations, the vortex being on and off the pole and differently distorted in each half of the month.

The standard deviation of the daily radiative heating distributions from the monthly-mean structure at η=6.7 in January 1979 (figure 5.16d) shows little variation ($<0.2 K day^{-1}$) equatorward of 20°N, reflecting the fact that the

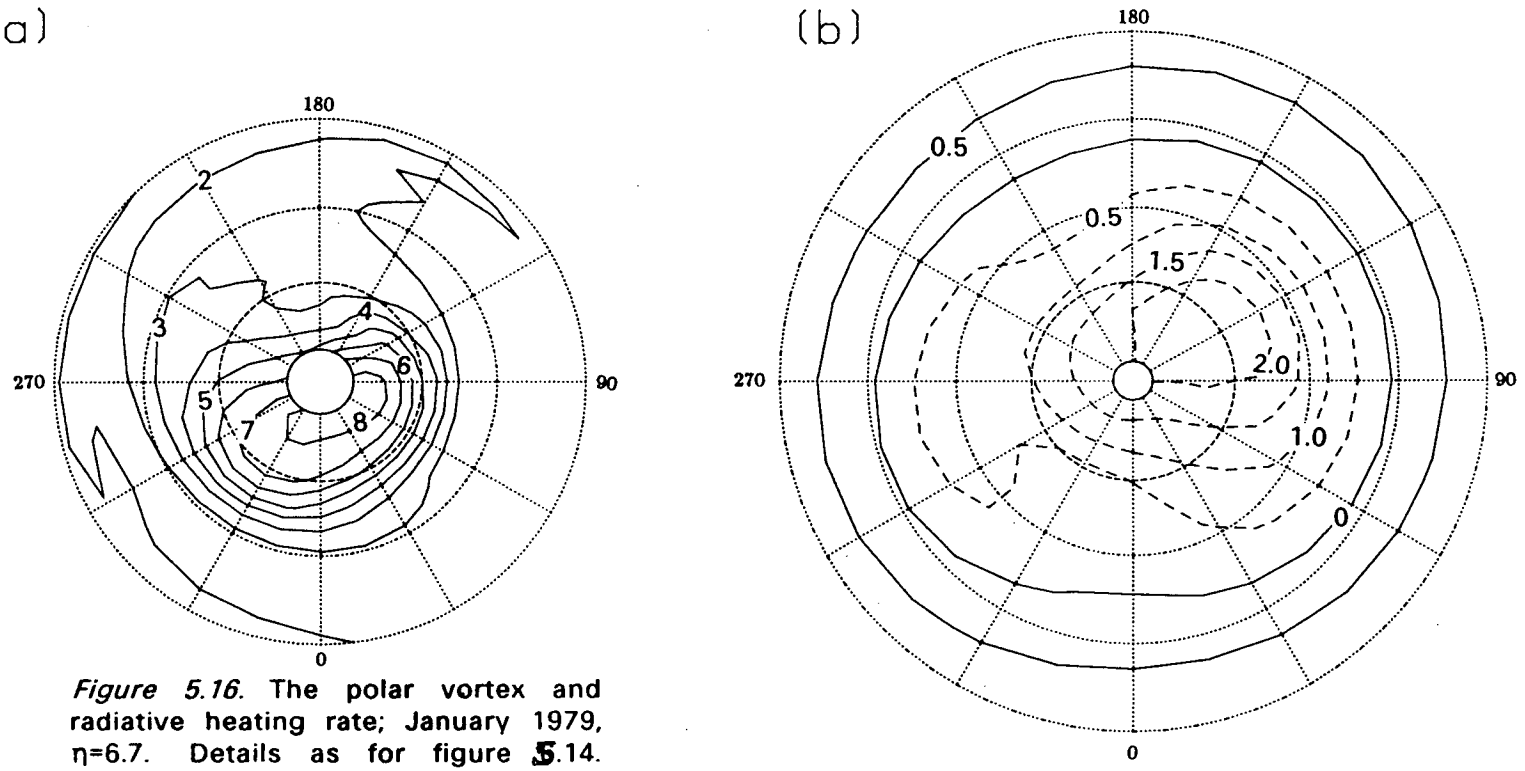
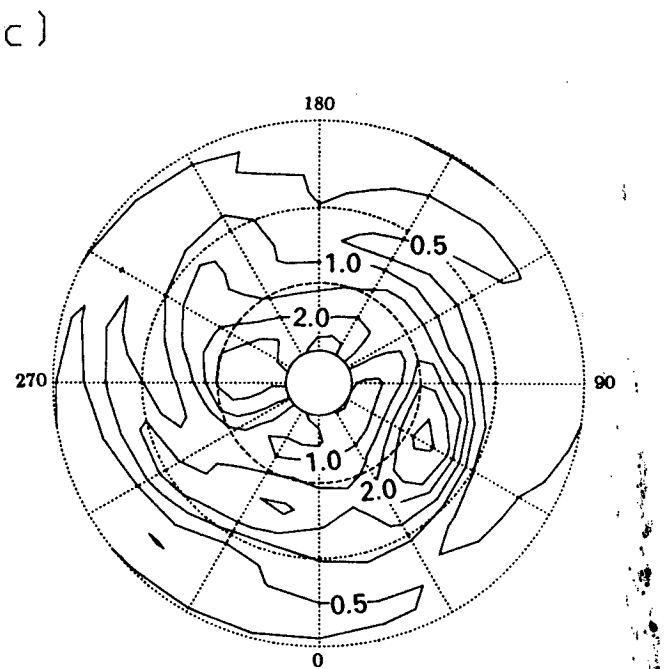


Figure 5.16. The polar vortex and radiative heating rate; January 1979, η=6.7. Details as for figure 5.14. (Figure continues.)



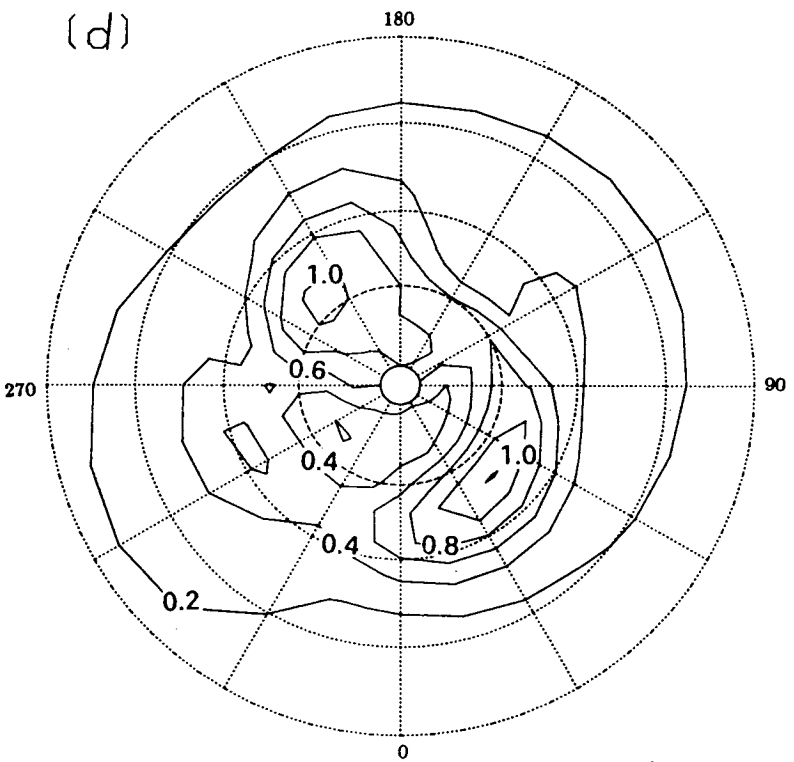


Figure 5.16 (continued). (c) standard deviation of the daily potential vorticity distribution about the monthly-mean field (contour interval $0.5\times10^{-6}kg^{-1}m^2s^{-1}$); (d) standard deviation of the daily heating distributions about the monthly-mean field (contour interval $0.2Kday^{-1}$).

temperature perturbations are of small amplitude and quite steady in the tropics. There are two pronounced maxima in the distribution of standard deviation, 50°E, 50°N and 220°E, 55°N, both of which exceed 1.2Kday^{-1} in regions where the net heating rate is 1-1.5Kday^{-1} (figure 5.16a); indeed, these two maxima are positioned to the East and West of the net cooling maximum, and represent oscillations in the position of this region through the month. A local minimum of standard deviation (<0.2Kday^{-1}) occurs 180° to the East of the January heating minimum, whilst a further local maximum (0.6Kday^{-1}) is to be found at 300°E, 40°N, where a tongue is evident in the net heating field.

d. February 1979, figure 5.17. The potential vorticity distribution shows that the vortex is smaller than in January 1979 and further displaced from the pole. The smaller size of the monthly-mean vortex arose not only because of erosion of material from its western flank, but also because of the large movements of the vortex in the middle of the month, after which it split (this will be considered in more detail later).

The heating distribution was much less distorted than in January; it was situated close to the pole in a predominantly wavenumber-2 configuration. The weakest cooling around each latitude circle occurred slightly to the western extreme of the vortex in the coolest air, whilst the strongest cooling once more occurred in the warm air to the East of the potential vorticity maximum. The region of net radiative warming, equatorward of around 36°N, showed very weak zonal asymmetries; the heating rate reaches 1Kday^{-1} at the equator.

Again, the daily evolution of the stratosphere is considered and used to interpret the standard deviation fields. After the warming near the end of January, the wavenumber-1 vortex re-established, persisting until 12 February, after which it became highly distorted and was forced off the pole as a secondary high-pressure region formed over the Pacific (17 February), with advection of low-potential vorticity air toward the polar cap, which caused the vortex to tear apart on 22 February; the whole system then persisted in this configuration whilst tongues of high potential vorticity air were eroded from each vortex.

It is the secondary vortex, formed after the sudden warming, which shows as a standard deviation of 1.5 kg^{-1}m^{2}s^{-1}s at 270°E, 55°N (figure 5.17c) in the

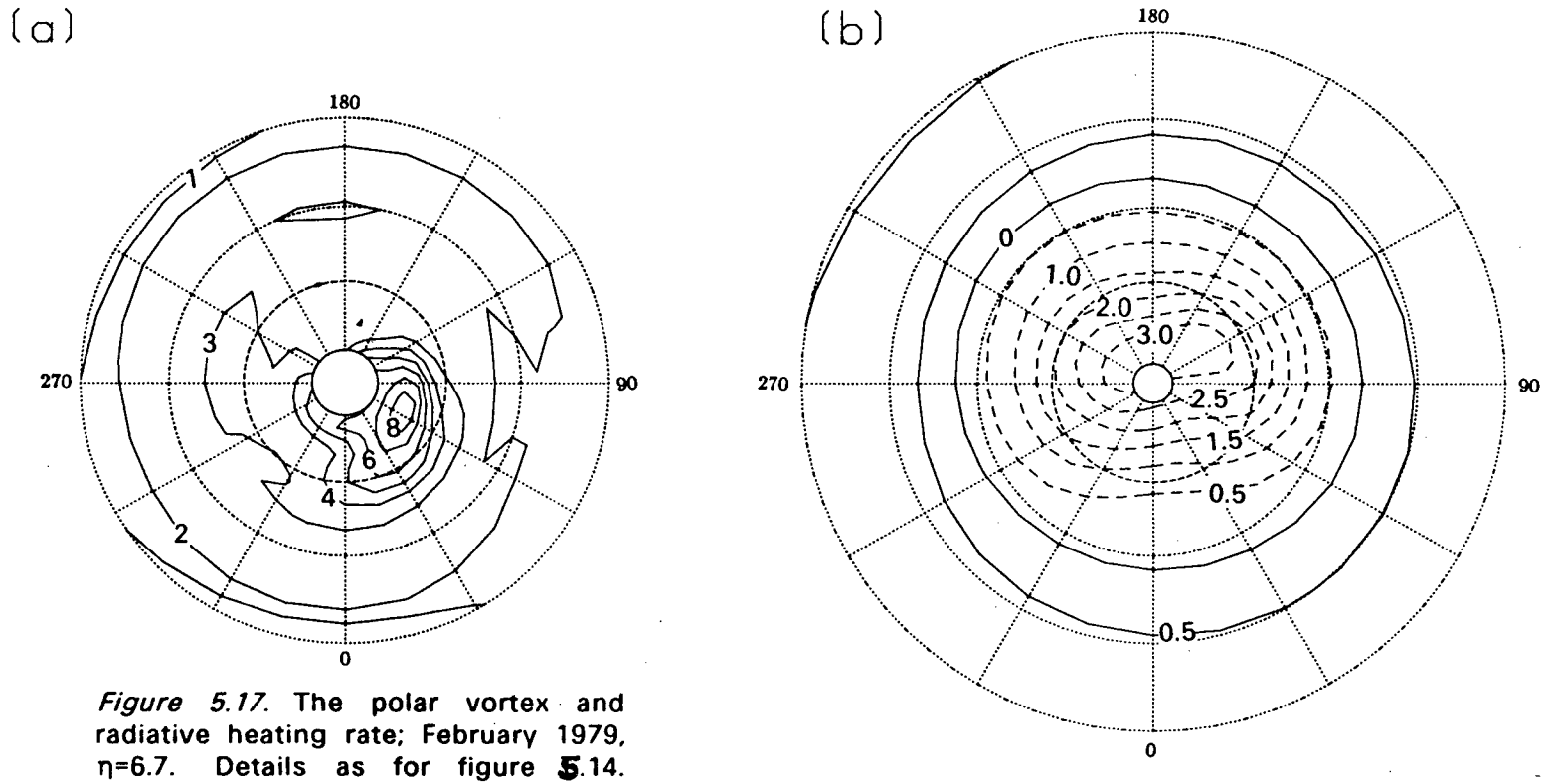
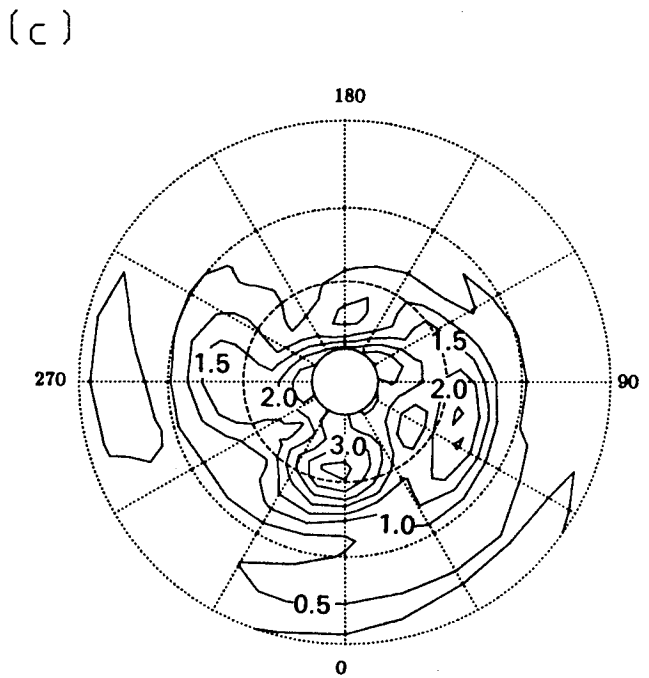


Figure 5.17. The polar vortex and radiative heating rate; February 1979, η=6.7. Details as for figure 5.14. (Figure continues.)



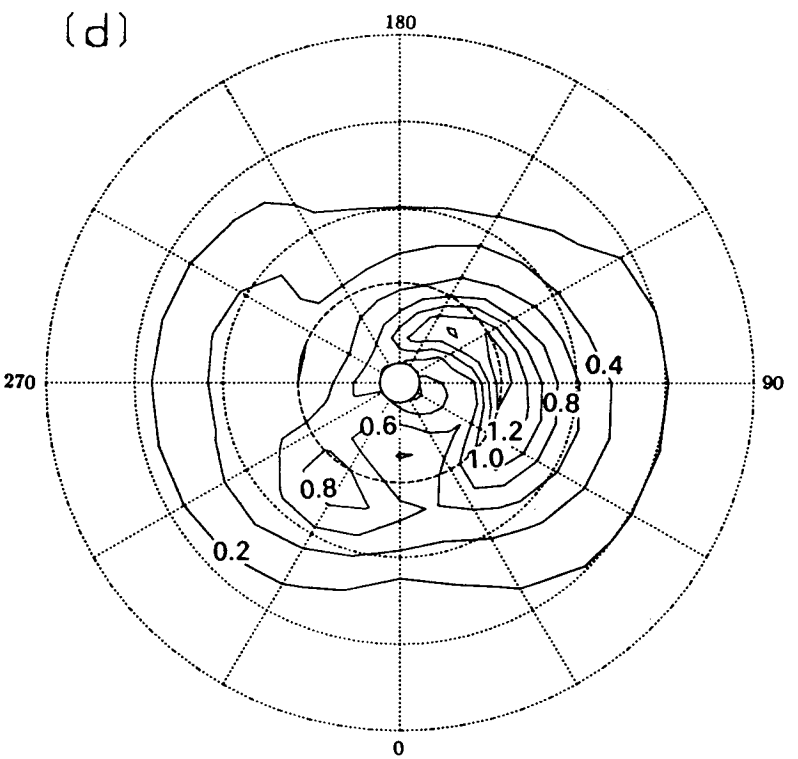


Figure 5.17 (continued). (c) standard deviation of the daily potential vorticity distribution about the monthly-mean field (contour interval $0.5\times10^{-6}kg^{-1}m^2s^{-1}$); (d) standard deviation of the daily heating distributions about the monthly-mean field (contour interval $0.2Kday^{-1}$).

potential vorticity field. The very strong deviation at 60°N on the Greenwich meridian arises because at the beginning of February this location was inside the polar vortex, whilst after the vortex split, low potential vorticity tropical air was advected across this region. Other differences arise because of differences in the vortex position, including the strong region near 55°N at 60–90°E.

The standard deviation of heating rate (figure 5.17d) also has a maximum in this region, associated with movements of warm air to the west of the cyclonic centre. It also shows variations resulting from heating differences as the equatorial air was advected across the pole in place of that in the vortex. The maximum variability of the heating rates is evident at 330°E, 55°N ($0.8 K day^{-1}$) and also in a strong feature centred at 120°E, 60°N, where it reaches $1.5 K day^{-1}$, which is 50% of the magnitude of the net cooling at this location, this being the strongest cooling observed in February. Stronger variations are evident at all latitudes than in January.

e. March 1979, figure 5.18. The vortex was by now very weak and the heating distribution showed only weak zonal-asymmetry poleward of 60°N and equatorward of 20°N; there was a fairly unstructured appearance to the field between these latitudes.

f. April 1979, figure 5.19. Both the potential vorticity and heating distributions were unstructured poleward of 20°N; generally there was weak heating equatorward and cooling poleward of around 60°N.

g. May 1979, figure 5.20. The zonality of the potential vorticity field shows that there was very little eddy activity; the heating distribution reflects this, having very little structure.

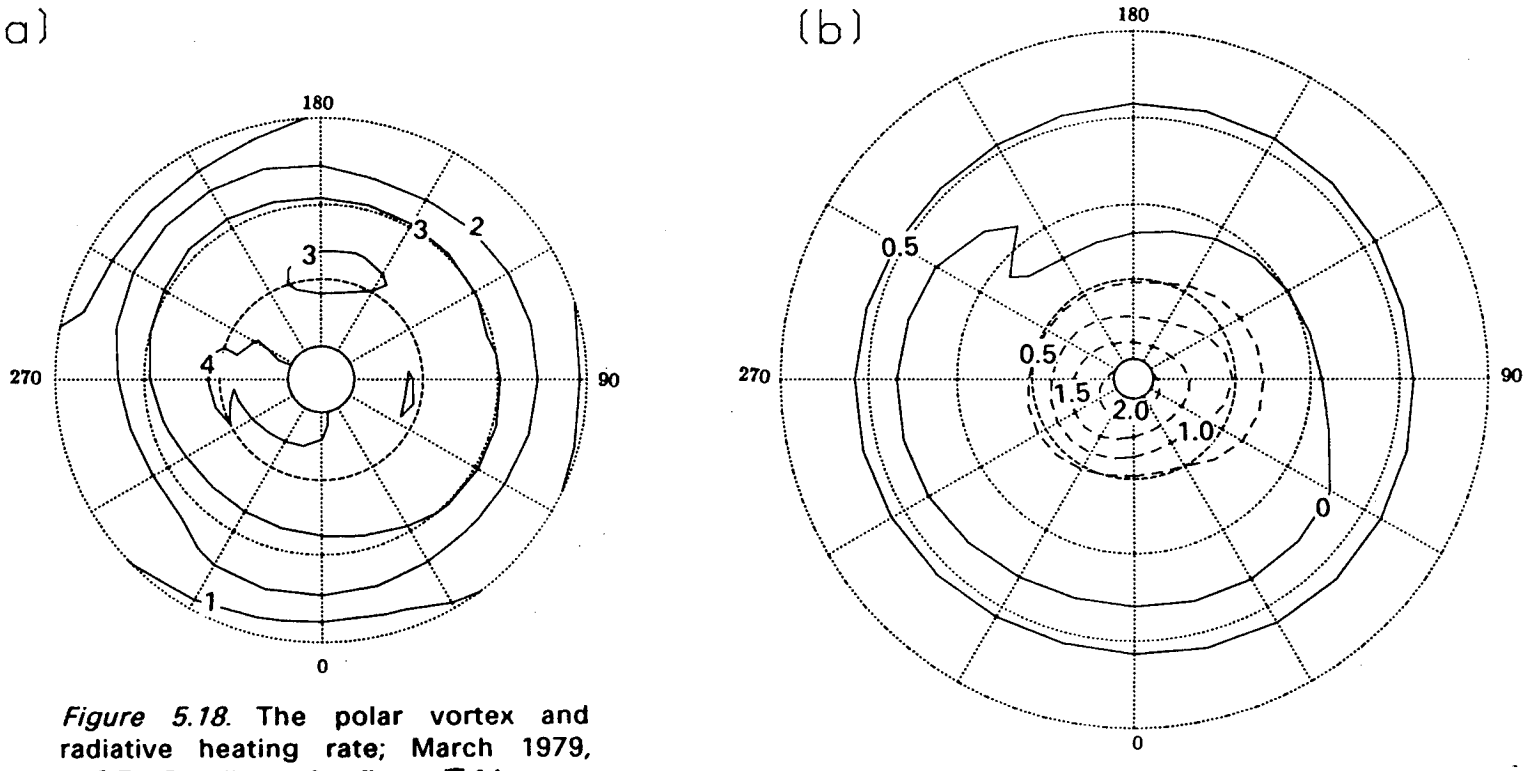


Figure 5.18. The polar vortex and radiative heating rate; March 1979, η=6.7. Details as for figure 5.14.

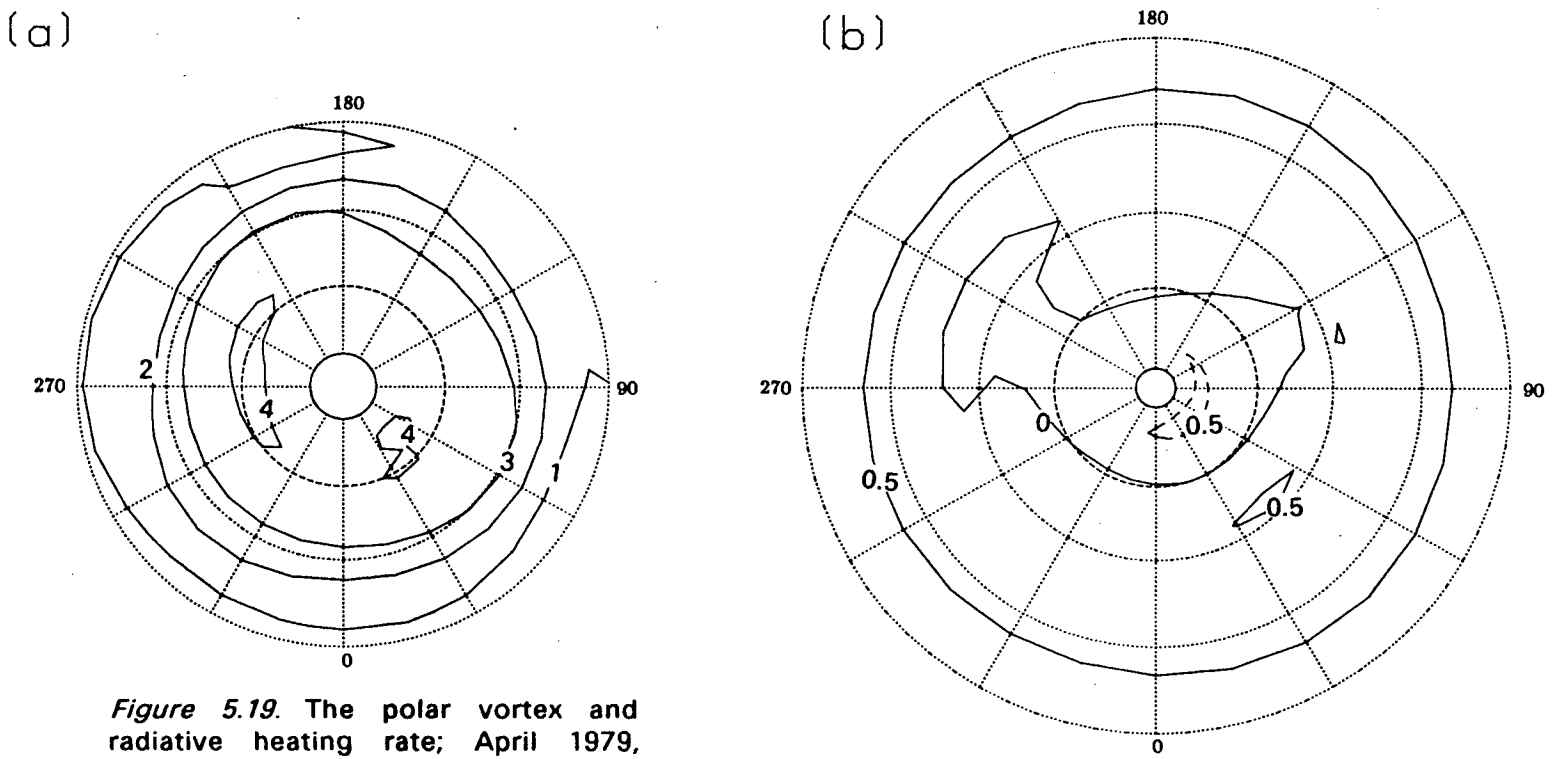


Figure 5.19. The polar vortex and radiative heating rate; April 1979, η=6.7. Details as for figure 5.14.

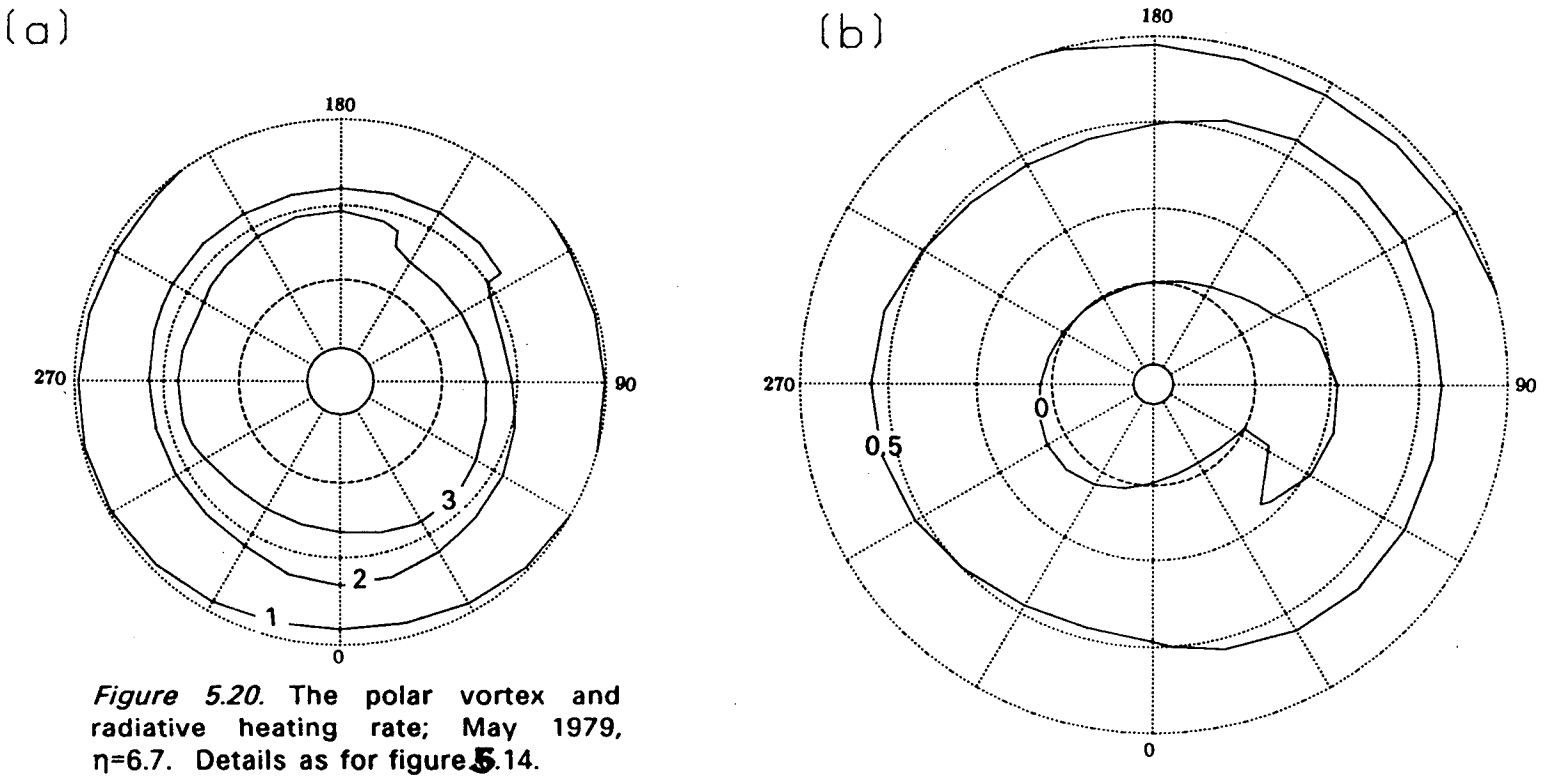


Figure 5.20. The polar vortex and radiative heating rate; May 1979, η=6.7. Details as for figure 5.14.

CHAPTER 6

APPLICATION TO ZONALLY-AVERAGED MODELS OF TRACER TRANSPORT.

6.1. The Philosophy of Zonal-Mean Modelling.

The great computational expense involved in calculating dynamical, radiative and photochemical processes interactively in numerical models of the atmosphere has necessitated the introduction of various approximations into the models of these processes. A good example is the use of random band models of the radiation transfer in preference to line-by-line calculations, as discussed in Chapter 4 of this thesis. Similarly, the dynamical processes may be parameterised.

The crudest parameterisation of the dynamics occurs in one-dimensional models, which consider a single column of the atmosphere, treating radiative and photochemical processes in some detail but including extremely crude representations of the dynamics. In such models the transport of trace gases is represented by some diffusion parameter. Because the transport in the real atmosphere varies greatly with location and season such models must represent only the crudest possible motions, and are perhaps valid only in horizontally homogeneous regions of the atmosphere; for instance, their application to an atmospheric column within the polar night vortex, a crude representation of which has been shown to be impermeable to mixing from outside by Juckes and McIntyre (1987), may have some value, whereas their application in the highly nonlinear surf-zone outside the vortex (McIntyre and Palmer, 1983, 1984) may be totally unjustified.

A more realistic parameterisation of the dynamical transport is possible within the zonally-averaged framework. If we consider linear, quasi-adiabatic perturbations on a slowly varying mean zonal state there is a theory which can explain the transport of trace gases by a mean-meridional-circulation and by the eddies. The isentropic coordinate formulation was developed Tung (1982) following the pioneering work of Reed and German (1965). The essential basis of the theory is that the eddy transport may be parameterised in terms of the zonal-mean tracer distributions so that the problem can be mathematically closed. The zonal-mean tracer transport equations will be derived in section 2 of this chapter, where more details of the approximations involved will be considered.

Despite the fact that the zonally-asymmetric motions observed in the atmosphere are not linear such zonally-averaged models have been applied to the problem of predicting monthly-mean tracer distributions with some success. Indeed, much of our current understanding of the distribution of trace gases and their evolution in the presence of pollutants caused by anthropogenic activity has come from the use of such models. It is the long-term mean behaviour of the atmosphere which can be predicted by these models, thus the parameterisation of the eddies must be regarded only as some statistically valid exercise rather than a description of the day-to-day transport in the stratosphere; for example, we cannot expect the daily evolution of the stratospheric state through a strongly nonlinear event (a sudden warming) to be reproduced. A consistent formulation thus requires only the slow, seasonal changes in the mean-meridional-circulation to be examined; we shall show that this necessitates only a consideration of the diabatic circulation, since the transient components, although significant on a daily basis, fluctuates throughout the month and has a negligible effect on tracer distributions over long time periods.

Many of the advantages and disadvantages of zonal-mean models of the middle atmosphere have been discussed by Harwood (1980), so they will not be repeated here. The emphasis of this chapter is to examine the role of the diabatic circulation in the evolving tracer distribution in the stratosphere - the eddy-transport is not discussed. Following the formulation of the theory in section 2, the remainder of the chapter is devoted to diagnosis of various terms in the tracer budget equation and a discussion of their relative importance. Section 3 contains a review of previous diagnoses of the diabatic circulation for comparison with our results in section 4. In section 5 the diabatic circulation is used to determine the cross-isentropic transport of trace gases; the mean transport is compared to the cross-isentropic eddy transport. In section 6 the implications of the results of chapter 5 are discussed and the possibility of developing a zonal-mean tracer transport model in isentropic coordinates discussed.

6.2. Formulation of the Zonally-Averaged Tracer Transport.

6.2.1. The Tracer Budget Equation.

For any trace constituent with mixing ratio μ and source term S, the budget equation is:

$$\frac{\partial}{\partial t}(\sigma_\eta \mu) + \nabla_\eta \cdot (\sigma_\eta \mu \mathbf{u}) + \frac{\partial}{\partial \eta}(\sigma_\eta \mu w) = \sigma_\eta S. \tag{6.1}$$

This equation states that the local change in (density-weighted) tracer amount is given by the three-dimensional convergence of its flux as well as the source term, which may be photochemical or due to surface release of the constituent.

The zonal-average of the trace gas distribution may be obtained using the identities (2.22) and (2.26) with the zonal-mean of equation (6.1) and simplified using the zonal-mean continuity equation (2.28), giving:

$$\begin{aligned}
&\frac{\partial}{\partial t}\langle\mu\rangle + \frac{\langle v\rangle}{a}\frac{\partial}{\partial \phi}\langle\mu\rangle + \langle w\rangle\frac{\partial}{\partial \eta}\langle\mu\rangle = \langle S\rangle \\
&-(\sigma_\eta)^{-1}\left\{\frac{1}{a\cos\phi}\frac{\partial}{\partial \phi}([\sigma_\eta \mu^* v^*]\cos\phi) - \frac{\partial}{\partial \eta}([\sigma_\eta \mu^* w^*])\right\} \\
&\quad = \langle S\rangle - ([\sigma_\eta])^{-1}\nabla_M \cdot \mathbf{\Lambda},
\end{aligned} \tag{6.2}$$

where $\mathbf{\Lambda} = (\Lambda_\phi, \Lambda_\eta) = ([\sigma_\eta v^* \mu^*], [\sigma_\eta w^* \mu^*])$. Hence, changes in the mass-weighted zonal-mean tracer distribution are brought about by zonal-mean advection, the zonal-mean source term and the divergent eddy fluxes of tracer in the meridional plane.

The zonal-mean formulation may also be performed using the conventional zonal average; the budget is then expressed by:

$$\begin{aligned}
&\frac{\partial}{\partial t}[\mu] + \frac{[v]}{a}\frac{\partial}{\partial \phi}([\mu]) + [w]\frac{\partial}{\partial \eta}([\mu]) = [S] \\
&+ ([\sigma_\eta])^{-1}\left([\sigma_\eta' S'] - \frac{\partial}{\partial t}([\sigma_\eta' \mu']) \right. \\
&\quad \left. - \frac{1}{a\cos\phi}\frac{\partial}{\partial \phi}([\sigma_\eta \mu' v']\cos\phi) - \frac{\partial}{\partial \eta}([\sigma_\eta \mu' w'])\right).
\end{aligned} \tag{6.3}$$

At first sight this formulation appears more complicated than that for the mass-weighted zonal-mean tracer distribution since it contains eddy source and transience terms; in practice these are hidden in the mass-weighted terms of equation (6.2) so that they are needed for a complete study with either formulation.

6.2.2. Parameterisation of the Eddy Tracer Fluxes.

The mass-weighted zonal-mean tracer transport equation (6.2) contains terms which require a knowledge of the eddy structure of trace constituent or some representation of it. In a zonal-mean models such fields are not known explicitly, so to close the problem mathematically some manner of parameterising these eddy fluxes must be developed. In this section a method of parametrisation in which the eddy fluxes are expressed in terms of the mean fields is reviewed: this is the K-theory, first introduced by Reed and German (1965), although it has been substantially revised as our understanding of the nature of the problem has developed. Although the derivation here is presented in isentropic coordinates there are analogous theories in other coordinate systems (most models have employed a logarithmic pressure coordinate system). The theory is outlined here so that some idea of the restrictions of zonal-mean modelling studies become apparent; an understanding of the limitations of the theory is essential to its successful application.

The K-theory has recently been considered by Tung (1982) and Plumb and Mahlman (1987); these studies provide the basis for the presentation of the theory in this thesis. The essence of the parametrisation is to express the eddy-flux of tracer in the meridional plane in terms of a matrix, K, and the mean tracer distribution, $<\mu>$, as:

$$\Lambda = -[\sigma_\eta]\underline{\underline{K}} \cdot \nabla_M <\mu>, \tag{6.4}$$

where

$$\underline{\underline{K}} = \begin{vmatrix} K_{\phi\phi} & K_{\phi\eta} \\ K_{\eta\phi} & K_{\eta\eta} \end{vmatrix}$$

This eddy transport tensor, K, may further be partitioned into antisymmetric, L,

and symmetric, D, parts:

$$\underline{\underline{K}} = \underline{\underline{L}} + \underline{\underline{D}} \tag{6.5}$$

The physical significance of such a partitioning is that the transport by the antisymmetric tensor, $\underline{\underline{L}}$, is advective, whilst that by the symmetric tensor, $\underline{\underline{D}}$, is diffusive in nature - this will become apparent as the precise form of the tensors is derived.

The starting point of the theory is the linearised perturbation equation for the eddy tracer distribution, obtained by subtracting equation (6.2) from equation (6.1) and neglecting all quadratic terms; this gives:

$$D_z\mu^* + \frac{v^*}{a}\frac{\partial}{\partial\phi}(<\mu>) + w^*\frac{\partial}{\partial\eta}(<\mu>) = S^*, \tag{6.6}$$

where $D_z \equiv \partial/\partial t + (a\cos\phi)^{-1}<u>\partial/\partial\lambda$ is the derivative following the zonal-mean zonal flow. The small amplitude assumption has already been introduced; this is essential in the neglect of the quadratic eddy terms and the advection by the mean-meridional-circulation in equation (6.6). The next step is to introduce the displacement fields, ξ_ϕ, ξ_η and ξ_s, related to the perturbation velocity and photochemical source terms by:

$$D_z(\xi_\phi, \xi_\eta, \xi_s) = (v^*\cos\phi, w^*, S^*); \tag{6.7}$$

thus ξ_ϕ and ξ_η define the northward and vertical displacement 'distances' from their mass-weighted zonal-mean locations. The assumption that the eddy photochemical source term may be expressed in terms of a linear deviation from its mass-weighted zonal-mean has also been made; the physical reality and interpretation of this assumption is less clear.

Substitution of the relations (6.7) into the perturbation tracer equation (6.6) gives:

$$D_z\mu^* + \frac{1}{a}\frac{\partial}{\partial\phi}(<\mu>)\frac{D_z\xi_\phi}{\cos\phi} + \frac{\partial}{\partial\eta}(<\mu>)\,D_z\xi_\eta - D_z\xi_s = 0.$$

The second major assumption of this derivation, that temporal variations in the

mean field occur on a much longer timescale than the perturbations, must now be introduced. The above equation may then be integrated to give the eddy tracer distribution in terms of the displacement fields:

$$\mu^* = \xi_s - \frac{\xi_\phi}{a\cos\phi}\frac{\partial}{\partial\phi}(<\mu>) - \xi_\eta\frac{\partial}{\partial\eta}(<\mu>). \qquad (6.8)$$

It is now possible to expand the eddy flux terms in the zonal-mean tracer budget equation (6.2). Using equation (6.8) it is trivial to see that:

$$\Lambda_\phi = [\sigma_\eta\xi_s v^*] - \frac{[\sigma_\eta\xi_\phi v^*]}{a\cos\phi}\frac{\partial}{\partial\phi}(<\mu>) - [\sigma_\eta\xi_\eta v^*]\frac{\partial}{\partial\eta}(<\mu>).$$

Use of equation (6.7) shows that:

$$[\sigma_\eta\xi_\phi v^*]\cos\phi = [\sigma_\eta\xi_\phi D_z\xi_\phi] = [\sigma_\eta]\frac{\partial}{\partial t}\left(\xi_\phi^2/2\right)$$

$$[\sigma_\eta\xi_\eta v^*] = [\sigma_\eta\xi_\eta D_z\xi_\phi] = [\sigma_\eta]\frac{\partial}{\partial t}[\xi_\phi\xi_\eta] - [\sigma_\eta\xi_\phi w^*]$$

(where the relation $w^*=D_z\xi_\eta$ has been used). The northward eddy flux of tracer may then be expressed as:

$$\Lambda_\phi\cos\phi = [\sigma_\eta\xi_s v^*]\cos\phi - \frac{[\sigma_\eta]}{a\cos\phi}\frac{\partial}{\partial t}\left(\frac{[\xi_\phi^2]}{2}\right)\frac{\partial}{\partial\phi}(<\mu>)$$

$$- \left([\sigma_\eta]\frac{\partial}{\partial t}([\xi_\phi\xi_\eta]) - [\sigma_\eta\xi_\phi w^*]\right)\frac{\partial}{\partial\eta}(<\mu>). \qquad (6.9)$$

A similar relationship may be derived for the vertical eddy tracer flux:

$$\Lambda_\eta = [\sigma_\eta\xi_s w^*]$$

$$- \frac{[\sigma_\eta\xi_\phi w^*]}{a\cos\phi}\frac{\partial}{\partial\phi}(<\mu>) - [\sigma_\eta]\frac{\partial}{\partial t}\left(\frac{[\xi_\eta^2]}{2}\right)\frac{\partial}{\partial\eta}(<\mu>). \qquad (6.10)$$

Now, comparison of equations (6.4), (6.9) and (6.10) shows that:

$$K_{\phi\phi} = \frac{1}{2\cos\phi}\frac{\partial}{\partial t}([\xi_\phi^2])$$

$$K_{\phi\eta} = \frac{\partial}{\partial t}([\xi_\eta \xi_\phi]) - ([\sigma_\eta])^{-1}[\sigma_\eta \xi_\phi w^*]$$

$$K_{\eta\phi} = ([\sigma_\eta])^{-1}[\sigma_\eta \xi_\phi w^*] \qquad (6.11)$$

$$K_{\eta\eta} = \frac{1}{2}\frac{\partial}{\partial t}([\xi_\eta^2])$$

There is also an eddy-source term:

$$S_E = \frac{1}{a\cos\phi}\frac{\partial}{\partial \phi}([\sigma_\eta \xi_s v^*]\cos\phi) + \frac{\partial}{\partial \eta}([\sigma_\eta \xi_s w^*]). \qquad (6.12)$$

(This too may be expressed in tensor form, but it is not necessary to do so for this thesis.)

The symmetric and antisymmetric components may be separated out in equation (6.11) to give $\underline{\underline{D}}$ and $\underline{\underline{L}}$:

$$D_{\phi\phi} = \frac{1}{2\cos\phi}\frac{\partial}{\partial t}([\xi_\phi^2])$$

$$D_{\phi\eta} = \frac{1}{2}\frac{\partial}{\partial t}([\xi_\phi \xi_\eta]) = D_{\eta\phi}$$

$$D_{\eta\eta} = \frac{1}{2}\frac{\partial}{\partial t}([\xi_\eta^2]) \qquad (6.13)$$

$$L_{\phi\phi} = 0 = L_{\eta\eta}$$

$$L_{\phi\eta} = \frac{1}{2}\frac{\partial}{\partial t}([\xi_\phi \xi_\eta]) - ([\sigma_\eta])^{-1}[\sigma_\eta \xi_\phi w^*] = -L_{\eta\phi}.$$

Following Tung (1982) it is convenient to express the eddy advective velocity associated with the antisymmetric tensor as:

$$(\langle v_E \rangle, \langle w_E \rangle) = ([\sigma_\eta])^{-1}(\partial([\sigma_\eta]L_{\phi\eta})/\partial\eta, -a^{-1}\partial([\sigma_\eta]L_{\phi\eta})/\partial\phi)$$

so that we may write:

$$\frac{1}{a}\frac{\partial}{\partial\phi}\left([\sigma_\eta]L_{\phi\eta}\frac{\partial}{\partial\eta}<\mu>\right) + \frac{\partial}{\partial\eta}\left([\sigma_\eta]L_{\eta\phi}\frac{1}{a}\frac{\partial}{\partial\phi}<\mu>\right) =$$

$$[\sigma_\eta]\left(\frac{<v_E>}{a}\frac{\partial}{\partial\phi}<\mu> + <w_E>\frac{\partial}{\partial\eta}<\mu>\right). \qquad (6.14)$$

(noting that derivatives of the cross-terms cancel). The tracer transport equation (6.2) may then be written:

$$\frac{\partial}{\partial t}\left(<\mu>\right)+\frac{(<v>+<v_E>)}{a}\frac{\partial}{\partial\phi}\left(<\mu>\right) + (<w>+<w_E>)\frac{\partial}{\partial\eta}\left(<\mu>\right) =$$

$$<S> + \left([\sigma_\eta]\right)^{-1}\{S_E + \nabla_M\cdot\left([\sigma_\eta]D\cdot\nabla_M<\mu>\right)\}. \qquad (6.15)$$

Here the mean flow advection and eddy advection are expressed on the left hand side, whilst the transport due to the symmetric tensor (the diffusive component) is included with the source terms on the right hand side.

Note that although the foregoing derivation was presented in isentropic coordinates it is equally valid in other coordinate systems; logarithmic pressure coordinates are often used.

Plumb and Mahlman (1987) have conducted a series of numerical experiments in which the validity of such a parametrisation is assessed. Three-dimensional data generated by a numerical model were used to derive eddy transport coefficients; these were then used in a zonal-mean model to examine its capability of reproducing the zonal-mean tracer fields derived from the parent model. The zonal-mean model fairly successfully reproduced the climatological, zonal-mean fields of the three-dimensional model, thereby justifying the suitability of the K-theory for examining trace constituent distributions on long timescales - say the monthly-mean fields. The assumptions inherent in the derivation render the theory unsuitable for examining the short term response of the atmosphere to perturbations: we cannot hope to represent the daily evolution of the zonal-mean tracer distribution through a stratospheric sudden warming.

6.3. A Review of Previously Diagnosed Diabatic Circulations.

In chapter 5 several recent studies of the zonal-mean net radiative heating rate were reviewed. These have all been used to determine diabatic circulations in the stratosphere and low mesosphere. In this section an

overview of these diabatic circulations is presented; similarities and differences between the calculations are discussed so that the results of this thesis can be compared to them. Again, December data are reviewed whenever possible.

a. Gille et al. (1987), December 1978, figure 6.1. The diabatic circulations presented in this study are consistent with the net heating distributions of Gille and Lyjak (1986); since the manner of adjusting the heating rates involved solving for the diabatic circulation (see Chapter 5) the method of solution will not be reviewed here.

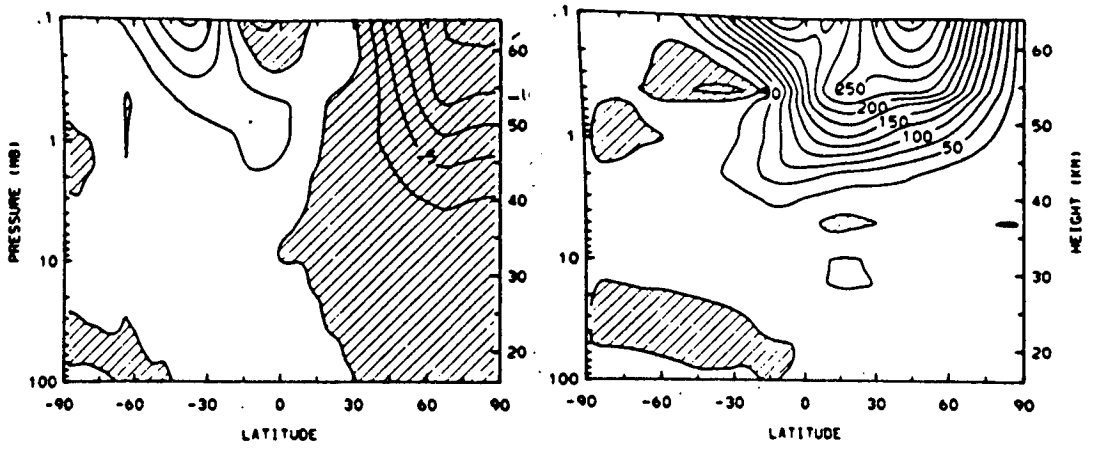


Figure 6.1. The zonal-mean diabatic circulation calculated by Gille *et al.* (1987) from the LIMS-derived net heating rate of Gille and Lyjak (1986). (a) the vertical velocity (mms^{-1}; contour interval $2.5mms^{-1}$); (b) the northward velocity (cms^{-1}; contour interval $2.5cms^{-1}$) for December 1978.

The diabatic circulation in December 1978 showed weak rising motion in most of the summer hemisphere and in the tropics with descent in the region of net radiative cooling in the northern hemisphere; the maximum descent rate was $12.5mms^{-1}$ above 0.2mbar (60km) at the North pole; at 0.5mbar (53km) the descent rate was slightly less than $10mms^{-1}$. The diabatic meridional velocity was very weak ($<0.25ms^{-1}$) at pressures greater than 3mbar, but generally northward, except for a small region in the low southern hemisphere. At pressures less than 3mbar ($z>40km$) there was significant northward flow

which exceeded 2.5ms^{-1} between 15-60°N at 55km and above.

b. Solomon et al. (1986), December 1978, figure 6.2. The heating rates of Kiehl and Solomon (1986) were used to determine the diabatic circulations in this study. An iterative method based on the transformed Eulerian-mean thermodynamic equation:

$$\frac{\partial}{\partial t}[T] + \frac{[v]^R}{a}\frac{\partial}{\partial \phi}[T] + [w]^R\left(\frac{HN^2}{R} + \frac{\partial}{\partial z}[T]\right) = \frac{[J]}{c_p}$$

was used. The iterative method initially assumed $[v]^R$ to be zero so that an estimate of $[w]^R$ could be obtained from the remaining terms, all of which were known. The $[w]^R$ field was adjusted so that its global mean vanished and then used to obtain an updated estimate of $[v]^R$ from the continuity equation. The iterative method proceeded by using the latest estimate of $[v]^R$. Five iterations were performed. Although Solomon *et al.* remark that the dominant term for determining $[w]^R$ is always the net radiative heating rate, we should note that there is a slight inconsistency between this method of obtaining the diabatic circulation and the method of adjusting the net heating rate, which did not allow for the northward mean heat transport.

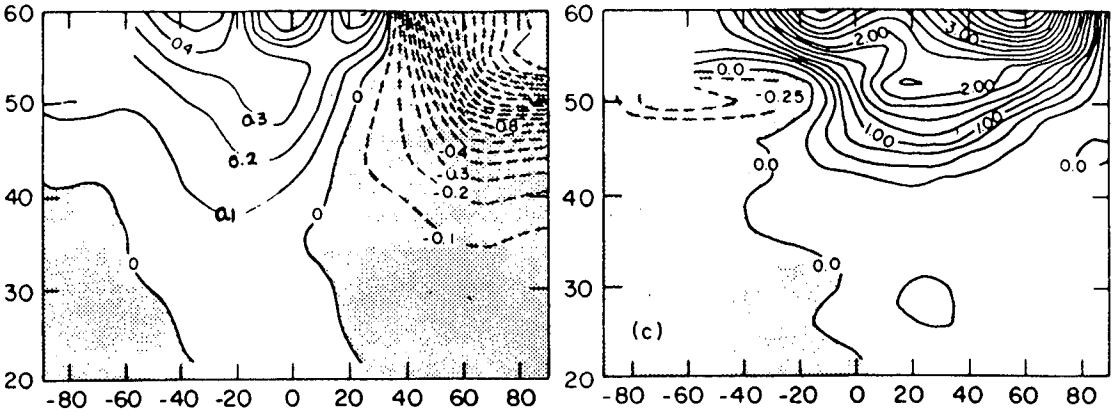


Figure 6.2. The zonal-mean diabatic circulation calculated by Solomon *et al.* (1986) using the net heating rates of Kiehl and Solomon (1986) calculated from LIMS data. (a) the vertical velocity (cms^{-1}, contour interval $0.1cms^{-1}$) for December 1978; (b) the northward velocity (ms^{-1}, contour interval $0.25ms^{-1}$) for January 1979.

The vertical motion field shows descent of $19ms^{-1}$ at 55km at the North pole, almost twice as strong as that of Gille *et al*, despite the fact that the North polar cooling rate calculated by Kiehl and Solomon is weaker than that of Gille and Lyjak. In equatorial regions there is better agreement between the two studies. Since the northward velocity field for December is not presented by Solomon *et al*, their January field is discussed; this shows weak flow below 40km but stronger northward velocities at higher levels: at 55km, 60°N the flow is faster than $3ms^{-1}$ and there is a strong northward jet at higher levels; such a strong northward flow is not obtained by Gille *et al.* for December or January.

c. Rosenfield et al. (1987), Climatological January, Figure 6.3. The diabatic circulation was calculated using the approximate, direct method of Dunkerton (1978), so that $[w]=[J]/c_p\Gamma$. The northward motion was thus obtained from the zonally-averaged continuity equation. Again there was a slight inconsistency between the method used to balance the heating rate and this calculation of the diabatic circulation; the global mean of [J] rather than $[J]/(\partial[\theta]/\partial z)$ was assumed to vanish and no account was taken of the northward heat transport.

January results are presented. There was weak ascent in the tropics,

reaching 2mms^{-1} at 0.5mbar (55km). The polar descent rate of 12mms^{-1} at the same level is slightly stronger than that of Gille *et al.*, but this could be explained by the different data sets and radiation models used. The northward velocity was negative in a shallow layer near 45km in the northern hemisphere; below this there was weak northward flow and at higher levels the northward motion was strong, exceeding 3ms^{-1} near 58km; this is slightly stronger than the result of Gille *et al.* but weaker than that of Solomon *et al.*

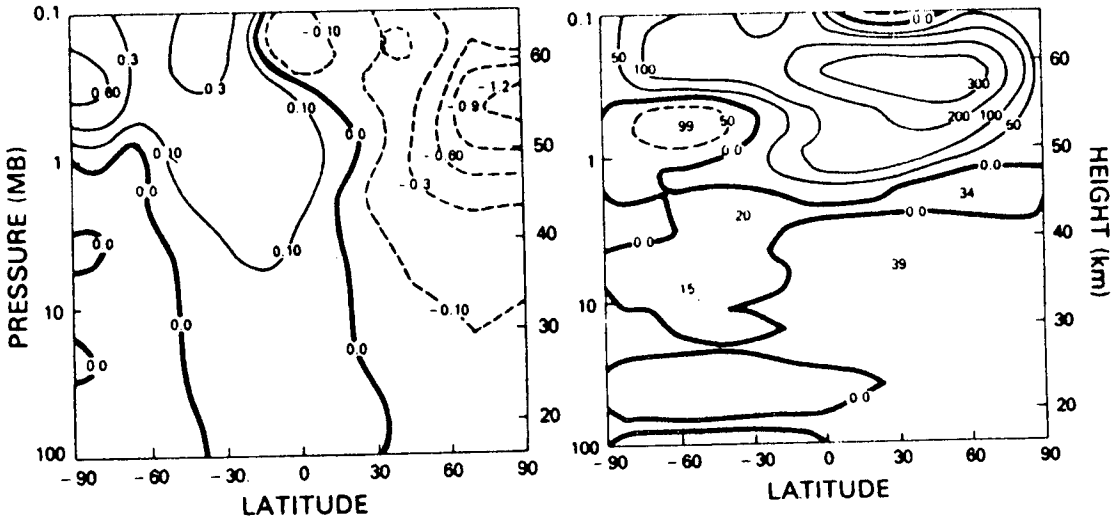


Figure 6.3. The diabatic circulation calculated from climatological January data by Rosenfield *et al.* (1987). (a) the vertical velocity (cms^{-1}, contour interval 0.1cms^{-1}); (b) the northward velocity (cms^{-1}, contour interval 50cms^{-1}).

d. Callis et al., Climatological and LIMS December, figure 6.4. The diabatic circulation calculated in this study is consistent with the adjustment of the radiative heating field since the approximation that $[W]=[J]/(c_p\Gamma_e)$ was used. The northward flow was obtained by assuming that it vanished at the South pole and integrating the steady-state continuity equation in the same manner as will be described in the next section. The vertical velocity in distance units was obtained from that in entropy units from the approximation $[dz/dt]=[\sigma_\eta][w]/\rho$.

Results show a descent rate of around 10mms^{-1} at θ=2500K (near 55km) at the North pole for the climatology (figure 6.4a) whilst the LIMS calculation showed a descent some 2mms^{-1} stronger (figure 6.4c), giving a value of 12mms^{-1}. Ironically, the climatological result agrees well with the LIMS result of Gille *et al.* whilst the LIMS result is more akin to the January climatological result of Rosenfield *et al.* The northward velocity reached 2ms^{-1} near 55km, 30-50°N for the climatology (figure 6.4b); the LIMS result being 1.2ms^{-1} weaker near 50°N, leaving a very weak 0.8ms^{-1} northward flow. The climatology appears to show a very strong southerly jet at the highest levels even in the northern hemisphere (although the contours are difficult to interpret); this jet is weaker in the northern hemisphere but stronger in the southern hemisphere in the LIMS calculations; comparison with the other studies leads one to question the reality of this jet.

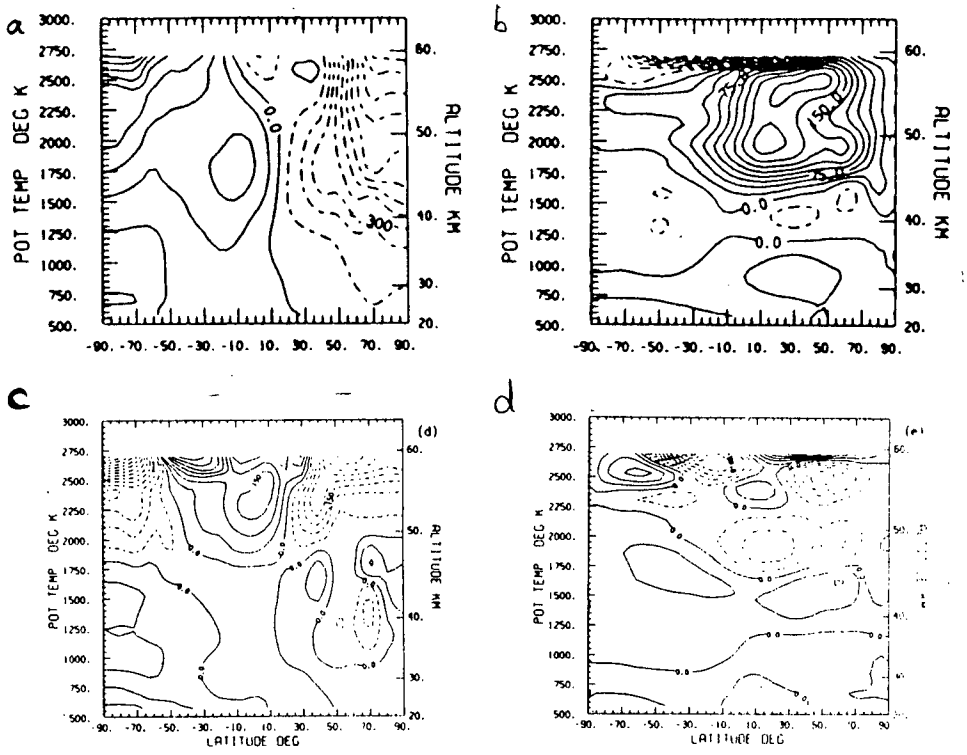


Figure 6.4. The diabatic circulation calculated by Callis *et al.* (1987) for December. (a) the climatological vertical velocity (kmday^{-1}, contour interval 0.1kmday^{-1}); (b) the corresponding northward velocity (kmday^{-1}, contour interval 2.5kmday^{-1}); (c) Difference between the LIMS and the climatological vertical velocity (positive region indicates that LIMS is stronger) (kmday^{-1}, contour interval 0.05kmday^{-1}); (d) as (c), but for the northward velocity (contour interval 0.3kmday^{-1}).

e. Overview. There is considerable variation in the magnitude, and even direction, of the diabatic circulation in the four studies; a brief summary is given in table 6.1. The descent rate near 55km at the North polar stratopause is close to 10mms^{-1} in all studies except that of Solomon *et al*, which shows 19mms^{-1}; this is curious, since their heating rate at the North polar stratopause shows weaker cooling than all of the other studies; moreover, the ascent rate near the tropical stratopause and in the southern hemisphere is in good agreement in all studies. Similarly there are discrepancies between the northward velocity fields; we must question the reality of the high-level jets calculated by Solomon *et al.* and Callis *et al.*: the former could result from the

upper boundary of the model domain being too low, but this is not a possible explanation in the latter case, where it must reflect a strong vertical gradient in the radiative heating rate, possibly caused by the radiation transfer model becoming invalid in the mesosphere; note that the radiation models of Gille *et al.* and Rosenfield *et al.* extend to higher levels than the other two and they do not calculate jets. The southward flow calculated Rosenfield *et al.* in the northern hemisphere mid-stratosphere is inconsistent with the other studies.

Note that of the pressure-coordinate studies only that of Gille *et al.* is entirely self-consistent, since the northward heat transport and the latitudinal variations in static stability were included in both the adjustment of the net radiative heating rate (so that the global-mean heating field need not vanish - see figure 5.1a) and the calculation of the diabatic circulation. The isentropic study is self-consistent, but does not allow the static stability to depart from its latitude-independent, radiative equilibrium value.

Table 6.1. Comparison of Aspects of the Zonal-Mean Diabatic Circulation in Previous Studies.

	A	B	C	D	E
Cooling Rate at winter polar stratopause ($Kday^{-1}$).	11	8	10	9	11
Corresponding descent rate (mms^{-1}).	12.5	19	12.5	10	13
Northward velocity at 20°N stratopause (ms^{-1}).	2.5	2	1.5	2	2.3

A=Gille *et al*; B=Solomon *et al*; C=Rosenfield *et al*; D=Callis *et al*, climatological; E=Callis *et al*, LIMS.

The variability of these results and the (sometimes unjustified) approximations used in obtaining them, suggest that a definitive study of the

zonal-mean diabatic circulation has not yet been made.

6.4. Diagnosis of the Diabatic Circulation in Isentropic Coordinates.

6.4.1. Method of Solution.

In this section the method of evaluating the northward motion from the globally-balanced vertical motion field is described. Denoting latitude by ϕ_j ($0<=j<=44$) (see figure 6.5) and assuming that the vertical flux is known at odd latitudes (ϕ_1, ϕ_3, ..., ϕ_{43}) then the northward mass-flux at even latitudes (ϕ_2, ϕ_4, ..., ϕ_{42}) may be obtained by integrating the finite difference form of the continuity equation:

$$\{[V_D]\}_j = \{[V_D]\}_{j-2} - \left(\partial\{[W]\}_{j-1}/\partial\eta\right)\left(\phi_j-\phi_{j-2}\right),\ j=2,4,6,...,44$$

(at each level k). The boundary condition applied was that $\{[V_D]\}_0=0$; i.e. that there is no northward flow through the southern boundary, which in this study was 88°S; this is sufficiently close to the South pole for the calculation to be sufficiently accurate and has the advantage that a regular latitudinal interval can be used without re-interpolation of the LIMS data, which is not the case if the pole is used as the boundary. Because of the global constraint applied to the vertical mass-flux we should also have $[V_D]_{44}=0$; in practice, numerical errors caused a very slight northward motion at 88°N, but this was insignificant. Reversing the direction of integration (i.e. setting $[V]_{44}=0$ and integrating southwards) gave changes of much less than 1% in the northward velocity at all latitudes.

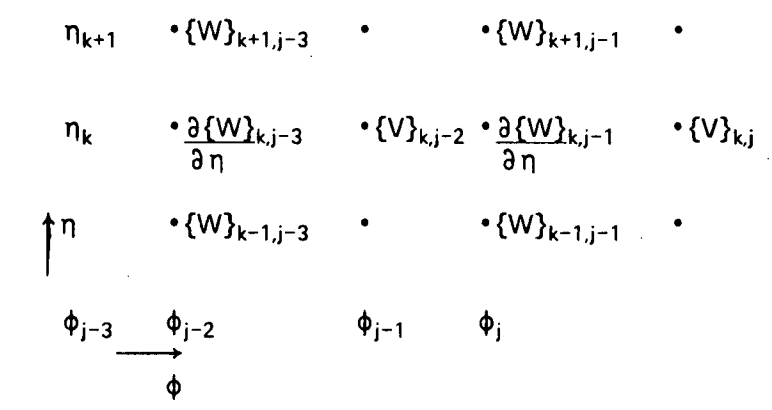


Figure 6.5. A portion of the latitude-height grid used in calculations of the zonal-mean diabatic circulation and tracer transport problems. The vertical mass-flux, $\{W\}_{k,j}$, is assumed to be known for odd j, so that $\partial\{W\}_{k,j-1}$ can be calculated; $\{V\}_{j,k}$ is then calculated at even latitudes.

From the mass-fluxes it is trivial to determine the mass-weighted zonal-mean velocities:

$$\{<v_D>\}_j = \{[V_D]\}_j/(\{[\sigma_\eta]\}_j\cos\phi_j) \qquad j=2,4,6,\ldots,42$$

$$\{<v_D>\}_j = 0, \qquad j=0,\ 44$$

$$\{<w>\}_j = \{[W]\}_j/\{[\sigma_\eta]\}_j \qquad j=1,3,5,\ldots 43.$$

To facilitate comparison with previous studies it is convenient to introduce an approximate vertical velocity, dz/dt. This has been calculated using the approximation of Holton (1986), where an approximate height, z, is introduced, where $dz=\kappa^{-1}Hd\eta=gN^{-2}d\eta$, where N is the buoyancy frequency. Assuming that $N^2\simeq 4\times10^{-4}s^{-2}$ this gives $dz\simeq 25\,d\eta$ (z in km). Distributions of vertical and northward velocities along with the 10-day travel distance are discussed in the next section.

6.4.2. Results Using LIMS Data in Isentropic Coordinates.

Because of the proportionality between vertical velocity and net radiative heating rate the pattern of vertical velocity is virtually identical to that of the heating fields discussed in chapter 5; in particular, the location of the zero-line is identical. The vertical and northward flow are discussed for each month of LIMS data; additionally, a vector plot of 10-day travel distance is presented, giving a visual impression of the strength of the meridional circulation.

a. November 1978, figure 6.6. The strongest descent rate, at the stratopause, was 12mms^{-1} near 76°N, η=7.8; there was descent throughout the northern hemisphere extratropics. There was also downward motion (relative to the isentropic surfaces) south of 60°S, which exceeded 1mms^{-1} between η=6.5-7.3. At higher levels the ascent was consistent with net radiative heating. The strongest upward motion was 4mms^{-1} near the South polar stratopause; there was also ascent in the tropics. The northward flow (figure 6.6b) was generally quite weak, exceeding 1ms^{-1} only in a small region of the lower mesosphere near 20°N; it was generally northward in the northern hemisphere, although there was a region of weak southerly flow in the tropics between η=7.0-7.3. The meridional flow was weak and southward throughout the southern hemisphere stratosphere.

b. December 1978, figure 6.7. The strongest descent rate near the North polar stratopause exceeded 12.5mms^{-1}, similar to that obtained by Callis *et al* using LIMS data, but slightly stronger than that of Gille *et al.* Our strongest ascent rate in the tropics exceeded 3mms^{-1} in the region where the heating rate was strongest to the South of the equator at η=7.6. The northward velocity was negative in most of the southern hemisphere, its magnitude exceeding 0.4mms^{-1} in middle latitudes below the stratopause. A region of southward flow extended through the northern hemisphere from the tropics to 60°N at η=7.0, similar to the result of Rosenfield *et al* but unlike the other calculations. As in the other recent studies the strongest northward flow occurred at the stratopause and in the low mesosphere from South of the equator to high northern latitudes; the strongest velocity of 1.8mms^{-1} at η=7.7 was slightly weaker than that of the other studies.

The ten day travel distance plot (figure 6.7c) gives a more visible impression of the strength of the meridional circulation and the distance the

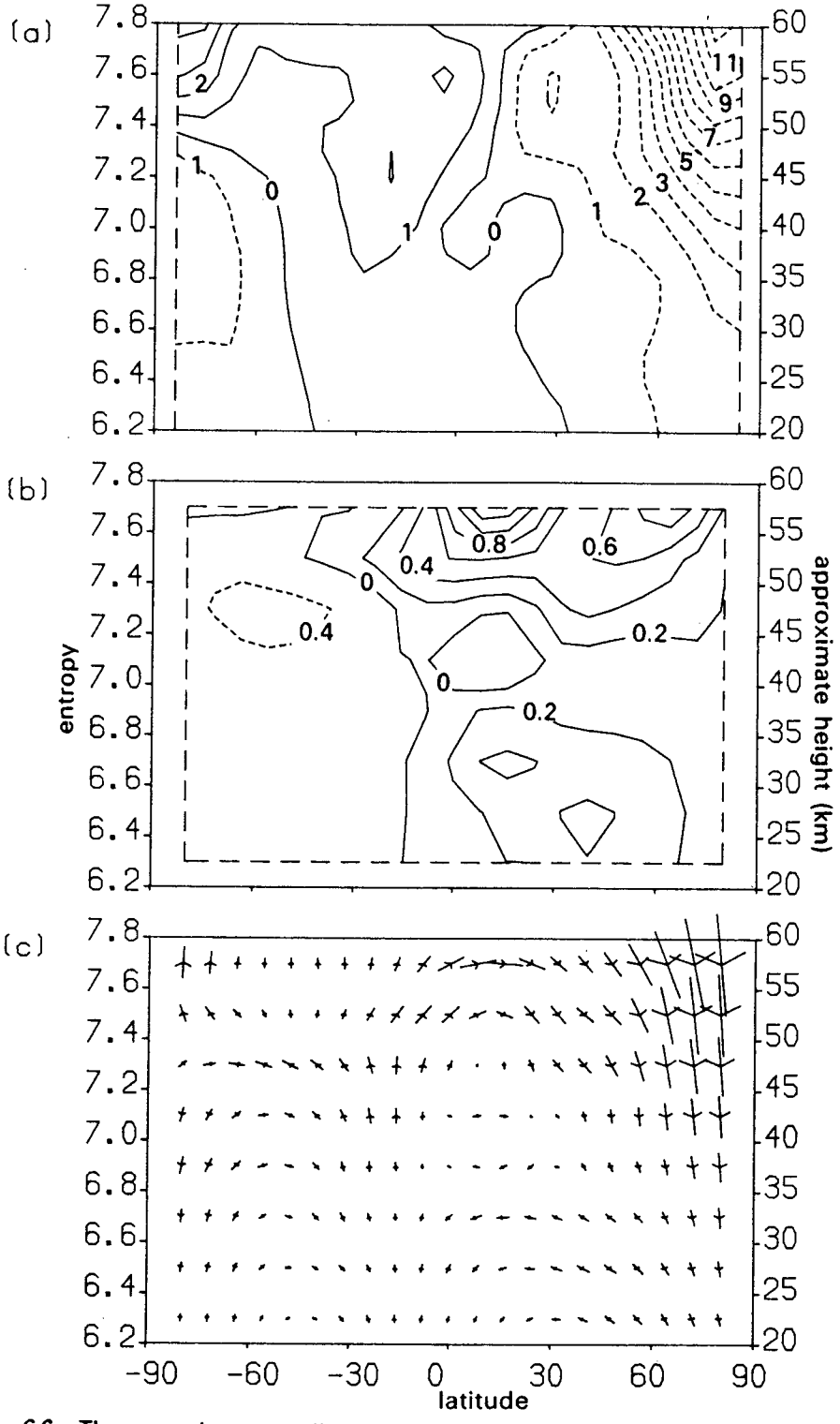


Figure 6.6. The zonal-mean diabatic circulation, November 1978. (a) Approximate vertical velocity, dz/dt (mms^{-1}), representing the rate of flow across isentropic surfaces (contour interval $1mms^{-1}$, ascent represented by solid lines, descent by dashed lines); (b) northward velocity (ms^{-1}, contour interval $0.2ms^{-1}$, solid lines represent northward flow, dashed lines represent southward flow); (c) the diabatic circulation represented vectorially, each arrow representing 10-day travel distance.

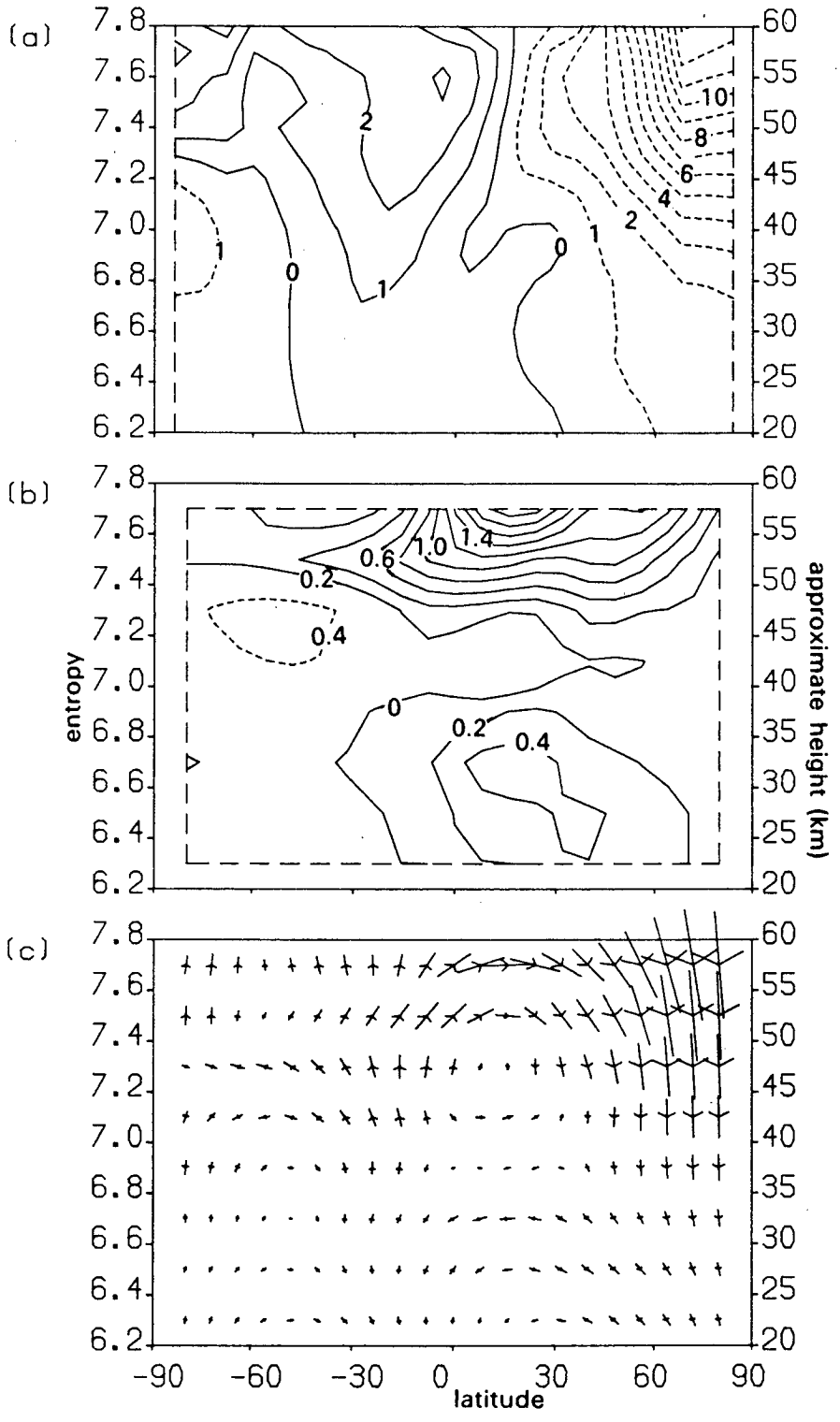


Figure 6.7. The zonal-mean diabatic circulation, December 1978. Details as for figure 6.6.

air parcels travel in that time. A parcel initially just above the stratopause in the polar night would be transported downwards to near the η=6.8 level in December (neglecting any mixing by isentropic eddies). Descent in the lower polar night is much weaker than this, as is the ascent rate in the tropics where the maximum vertical transport distance is around 0.2 units of entropy, or 10km, in a month. In the middle and low stratosphere (near η=6.7) the northward motion can transport air parcels 10° northward in December, whilst around the stratopause the northward transport is 15° of latitude in ten days. This strong northward transport from the tropics to middle and high latitudes of the winter hemisphere where the descent rate is strong is likely to play an important role in the transport of trace gases; although the timescale may be too long for significant transport of ozone from the tropics to the polar night, it must be important in determining the distribution of more inert trace gases. In the summer hemisphere the diabatic circulation is much weaker than in the winter, but is significant because of the lack of eddy transport there.

c. January 1979 figure 6.8. The vector plot (figure 6.8c) shows that the vertical travel distances were similar to those of December but that the northward flow was somewhat different. This is borne out by the plot of meridional velocity (figure 6.8b), which was northerly throughout the northern hemisphere and at low and high levels in the southern tropics. The jet exceeded $2ms^{-1}$ between 10-40°N at the highest level, which will cause enhanced transport of air parcels between the tropics and middle latitudes of the northern hemisphere. In the lower stratosphere the northward flow was almost identical to the previous month. In polar regions the vertical flow (figure 6.8a) was very similar to December, but at the equator there is descent at the highest level, consistent with the region of cooling associated with the downward propagation of the warm phase of the semi-annual oscillation. The strongest ascent, $3mms^{-1}$, was stronger than in December.

d. February 1979, figure 6.9. The decrease in net cooling at the winter polar stratopause led to weaker descent than in the previous months: the maximum downward velocity being reduced to $8mms^{-1}$ at the stratopause. At lower levels the decent rate is comparable to that of December and January. The maximum ascent rate in the tropics was slightly weaker than in January, exceeding $2mms^{-1}$ and was displaced southwards because of the radiative cooling associated with the downward propagation of the semi-annual

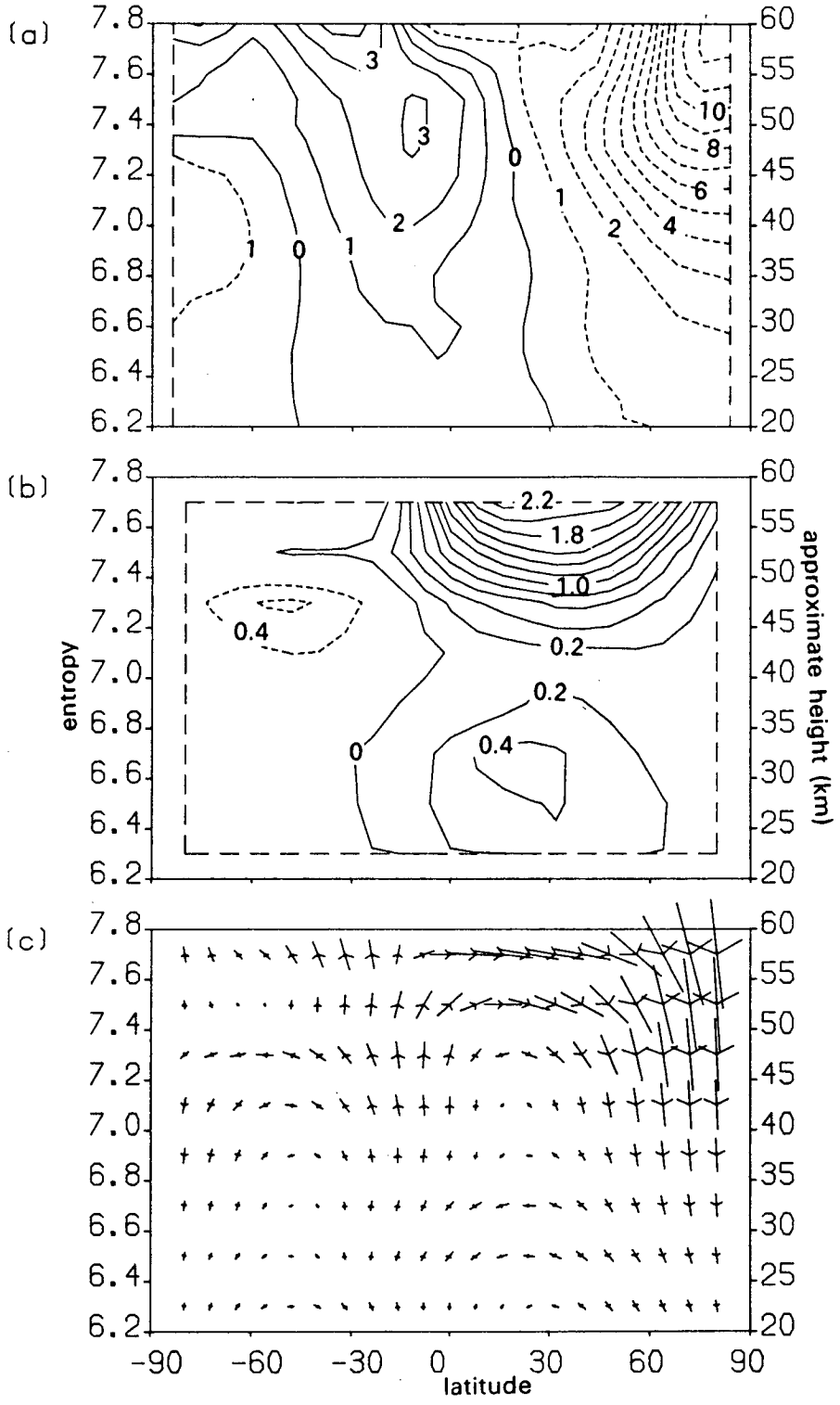


Figure 6.8. The zonal-mean diabatic circulation, January 1979. Details as for figure 6.6.

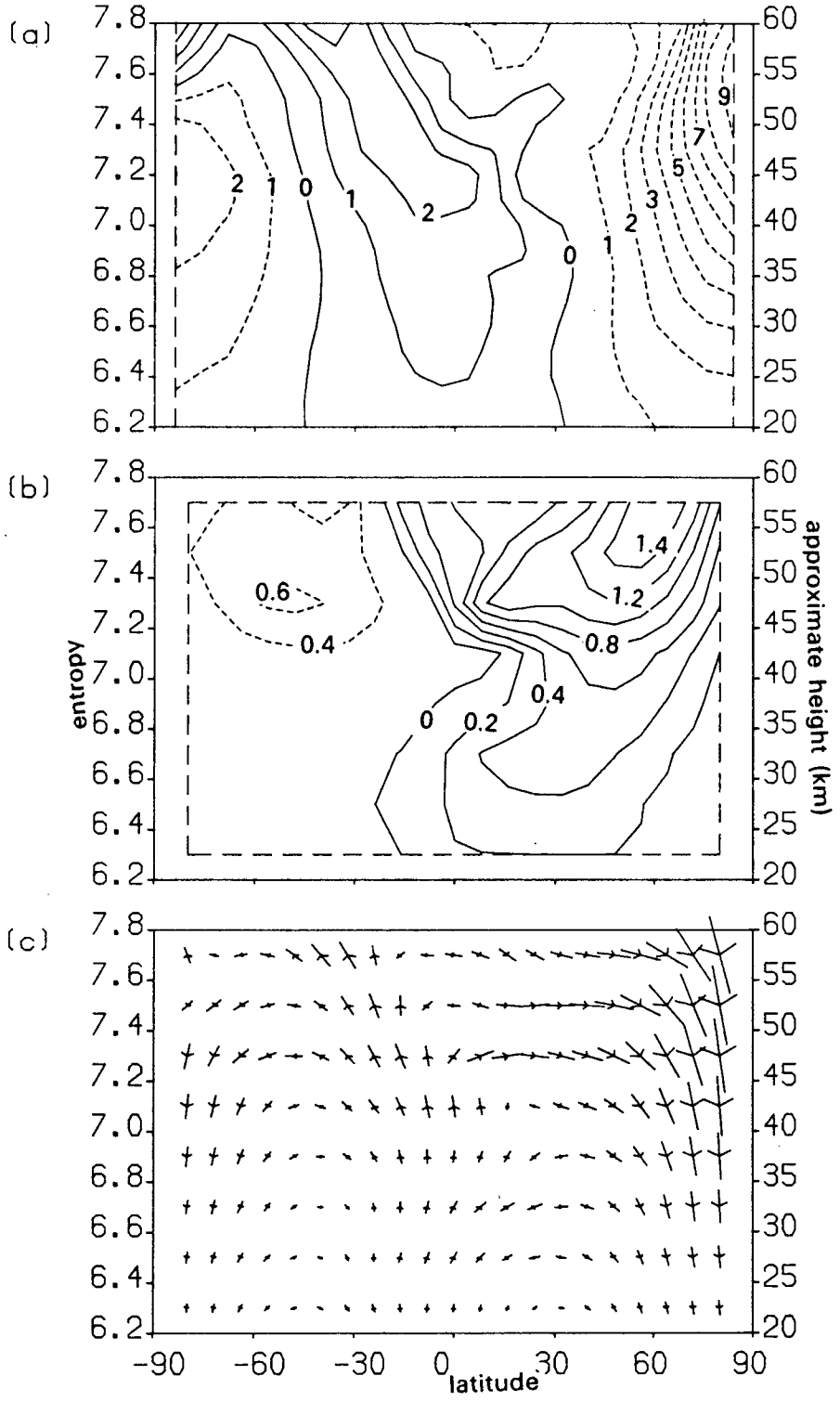


Figure 6.9. The zonal-mean diabatic circulation, February 1979. Details as for figure 6.6.

oscillation into the stratosphere. This high-level cooling caused a large change in the vertical mass-flux convergence in the tropics, so that the northward flow was much weaker at the tropical stratopause than in December. The strongest northward flow in February was $1.4ms^{-1}$ near 60°N. Northward flow is evident in most of the northern hemisphere, with southward flow, exceeding $0.6ms^{-1}$ near 60°S elsewhere.

e. March 1979, figure 6.10. The strong northerly jet linking the tropics to the middle latitudes near the stratopause vanished this month (figure 6.10c); there was weak northward flow in almost all of the northern hemisphere (with a maximum magnitude of $4ms^{-1}$), except above η=7.4 in the tropics where it was southwards. The southward flow exceeded $0.6ms^{-1}$ near the southern stratopause. In accord with the net radiative heating distribution the vertical velocity field was almost hemispherically symmetric (figure 6.10a), with descent of $5ms^{-1}$ at each pole and ascent between 30°S-30°N, except for a small region of descent at high levels at the equator.

f. April 1979, figure 6.11. The descent rate at the South pole reached $12mms^{-1}$ as it moved into its winter state; there was weak descent throughout the southern hemisphere tropics and weak ascent in almost all of the northern hemisphere except for the region poleward of 40-60°N below η=7.4, where there was weak descent. The strongest ascent in the northern hemisphere was $6mms^{-1}$ at η=7.6 at the pole. The meridional flow was weak and northward in most of the extratropical northern hemisphere and southward elsewhere, exceeding $0.6ms^{-1}$ near the stratopause.

g. May 1979, figure 6.12. The southward flow in the low southern hemisphere mesosphere was still quite weak, exceeding $0.6ms^{-1}$ near the equator and again near 60°S; the pattern was similar to (but weaker than) that of November 1978 (figure 6.6). The vertical motion showed a peak descent rate of $13mms^{-1}$ at the South polar stratopause and the ascent rate at the North pole was $5mms^{-1}$. There was weak ascent in the tropics.

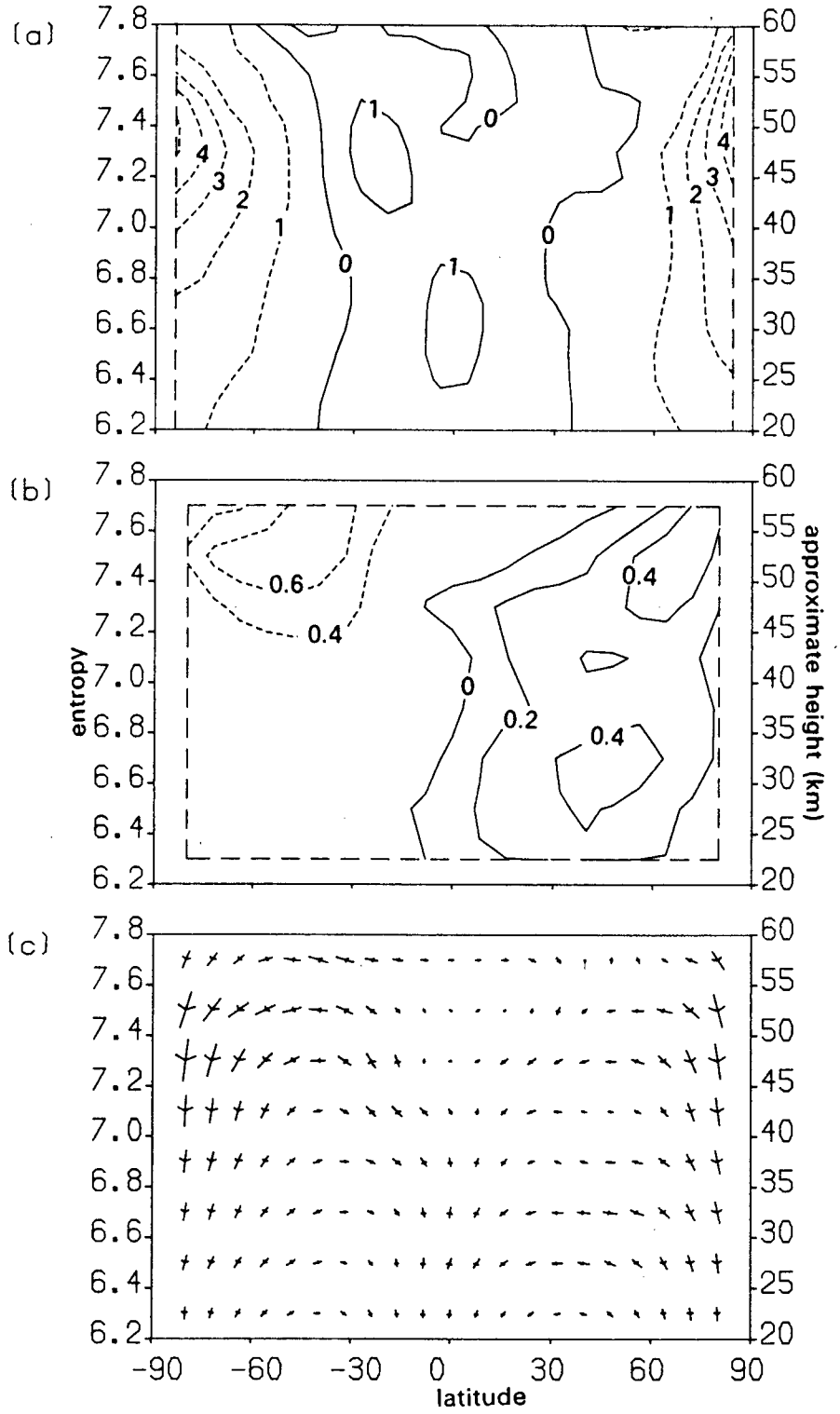


Figure 6.10. The zonal-mean diabatic circulation, March 1979. Details as for figure 6.6.

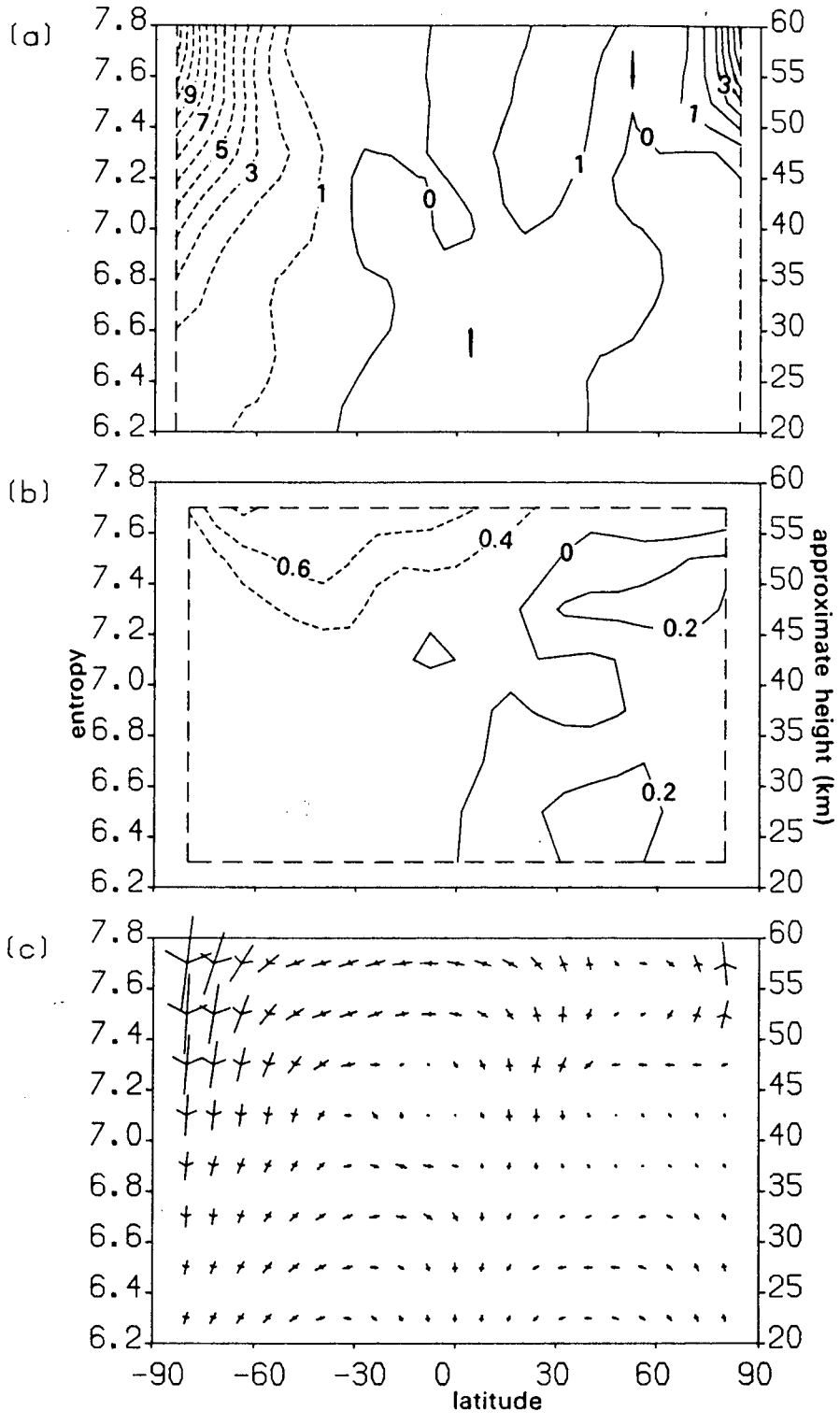


Figure 6.11. The zonal-mean diabatic circulation, April 1979. Details as for figure 6.6.

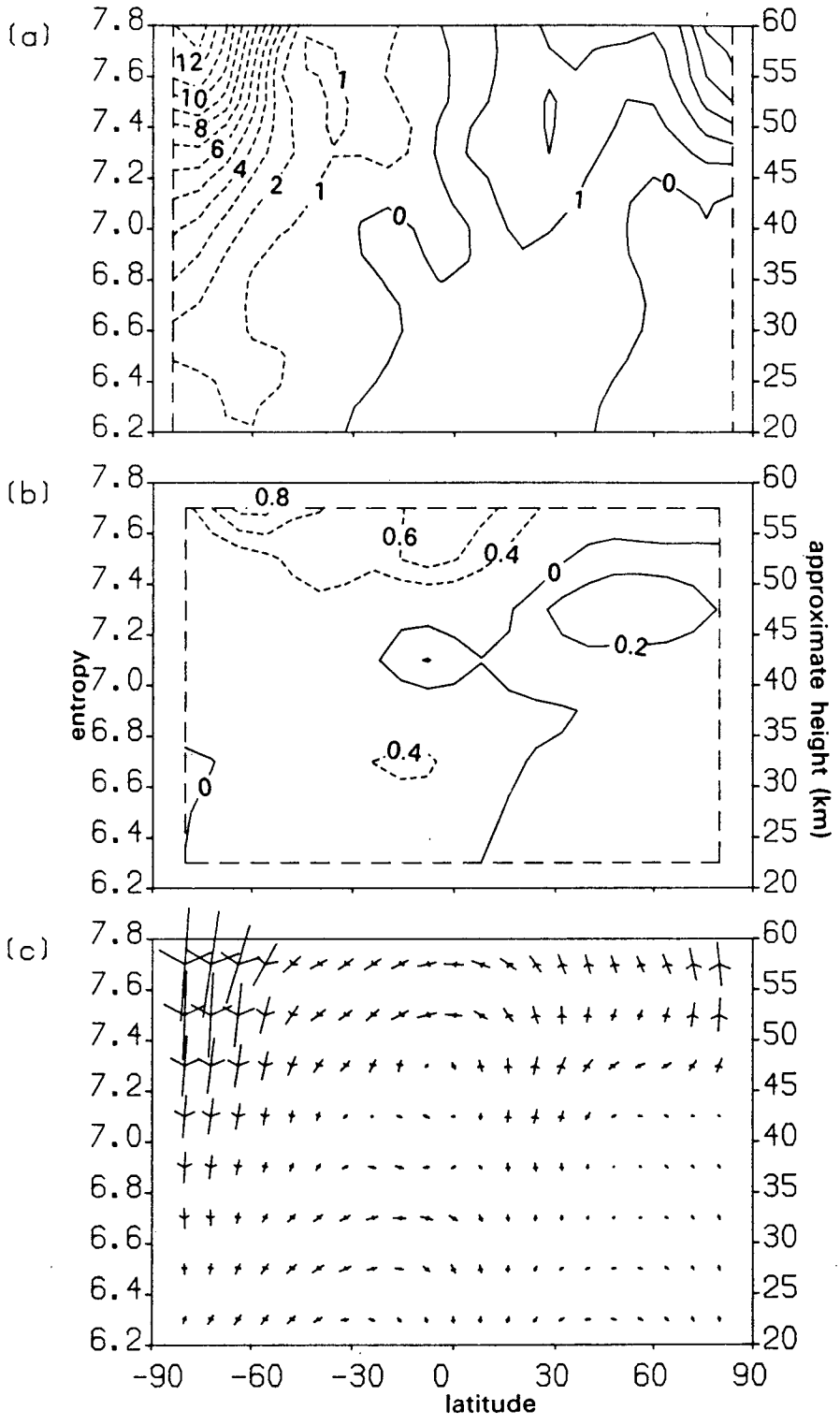


Figure 6.12. The zonal-mean diabatic circulation, May 1979. Details as for figure 6.6.

6.5. Zonally-Averaged, Cross-Isentropic Tracer Transport: Mean-ν-Eddy.

6.5.1. The Importance of the Problem.

In section 2 above the K-theory of tracer transport was outlined; this involves parameterising the eddy transport of trace gases in terms of the mean fields. Horizontal and vertical transport are included in the parameterisation, representing motion along isentropic layers and across isentropic layers, the latter component being due to diabatic eddies. Tung (1982) used a simple analytic approximation to propose that the cross-isentropic transport by diabatic eddies is negligible in comparison to the mean diabatic transport and the transport by isentropic eddy motions. In this section our data are used to provide observational estimates of the mean and eddy cross-isentropic transport of trace gases. The significance of this study is that if the transport due to diabatic eddies can be smaller than the cross-isentropic transport due to the mean diabatic circulation, it can be neglected in the parameterisation; that is, $K_{\eta\eta}$, $K_{\eta\phi}$ and $K_{\phi\eta}$ can be assumed to vanish. This gives a conceptually simple view of tracer transport in isentropic coordinates: it may be regarded as transport by a mean meridional circulation with an additional, isentropic transport due to eddies. Such a simplification is not possible in non-isentropic coordinate systems, since vertical motions there can occur in adiabatic flow: they are related to undulations of the isentropic surfaces with pressure or height.

By definition an adiabatic eddy has $w^* \equiv 0$, so that it has no cross-isentropic eddy transport. Tung (1982) showed that for a quasi-adiabatic eddy with a radiative dissipation time of 5 days, the cross isentropic eddy transport is negligible in comparison to the mean advection across the isentropic surface. Such an approximation is perhaps good in the middle stratosphere but at higher levels the radiative dissipation time is much shorter than this, so that the vertical eddy transport could be important. Further, the concept of using linearised radiative dissipation theory is an approximation which may not be applicable in the presence of large amplitude, nonlinear perturbations. From our knowledge of the non-zonal distribution of vertical velocity at isentropic levels we are able to make a direct deduction of the magnitudes of both mean and eddy transport so that a more accurate assessment of their relative roles may be given.

6.5.2. Method of Calculating the Transport.

The finite difference scheme used to calculate the cross-isentropic transport of trace gases by the mean and eddy motions is described here.

In finite difference form, the vertical gradient of any quantity s is expressed as:

$$\{\partial s/\partial \eta\}_k = \left(\{s\}_{k+1} - \{s\}_{k-1}\right)/\Delta\eta$$

where $\Delta\eta = \{\eta\}_{k+1} - \{\eta\}_{k-1}$. (Recall that k is the vertical index, and the vertical grid is that shown in figure 6.5.) Thus, the advection of trace gas by the vertical velocity is:

$$\{<w>\frac{\partial}{\partial\eta}\left(<\mu>\right)\}_k = \{<w>\}_k \frac{\left(\{<\mu>\}_{k+1} - \{<\mu>\}_{k-1}\right)}{\Delta\eta}$$

which is the product of the vertical gradient of mass-weighted zonal-mean trace mixing ratio with the vertical velocity.

The cross-isentropic eddy transport was obtained by substituting the eddy flux of tracer, $[\sigma_\eta w^*\mu^*]$, in place of s in the definition of vertical derivative above and dividing by the zonal-mean density.

6.5.3. Mean and Eddy Ozone Transport.

The ozone distribution at η=7.5 in December (figure 5.8a) shows a distinct zonally-asymmetric pattern between 20-60°N which leads to a local maximum in the net radiative heating rate (figure 5.7b), as discussed in chapter 5. The diabatic circulation in December (figure 6.7) shows strong descent at high levels in the polar cap, with weaker descent at lower levels and rising motion in the tropics, which is strongest to the south of the equator. The mass-weighted zonal-mean distribution of ozone (figure 6.13a) shows a maximum mixing ratio of almost 9ppmv near η=6.8, also just to the south of the equator, with a decrease in mixing ratio above and below this level; in mid-latitudes the peak mixing ratio occurs slightly higher than in the tropics (near η=7.0). The vertical gradient in ozone mixing ratio is thus positive in the lower stratosphere and negative in the upper stratosphere. Consequently, the mean vertical advection term, $<w>(\partial<\mu>/\partial\eta)$, is positive at low levels in the

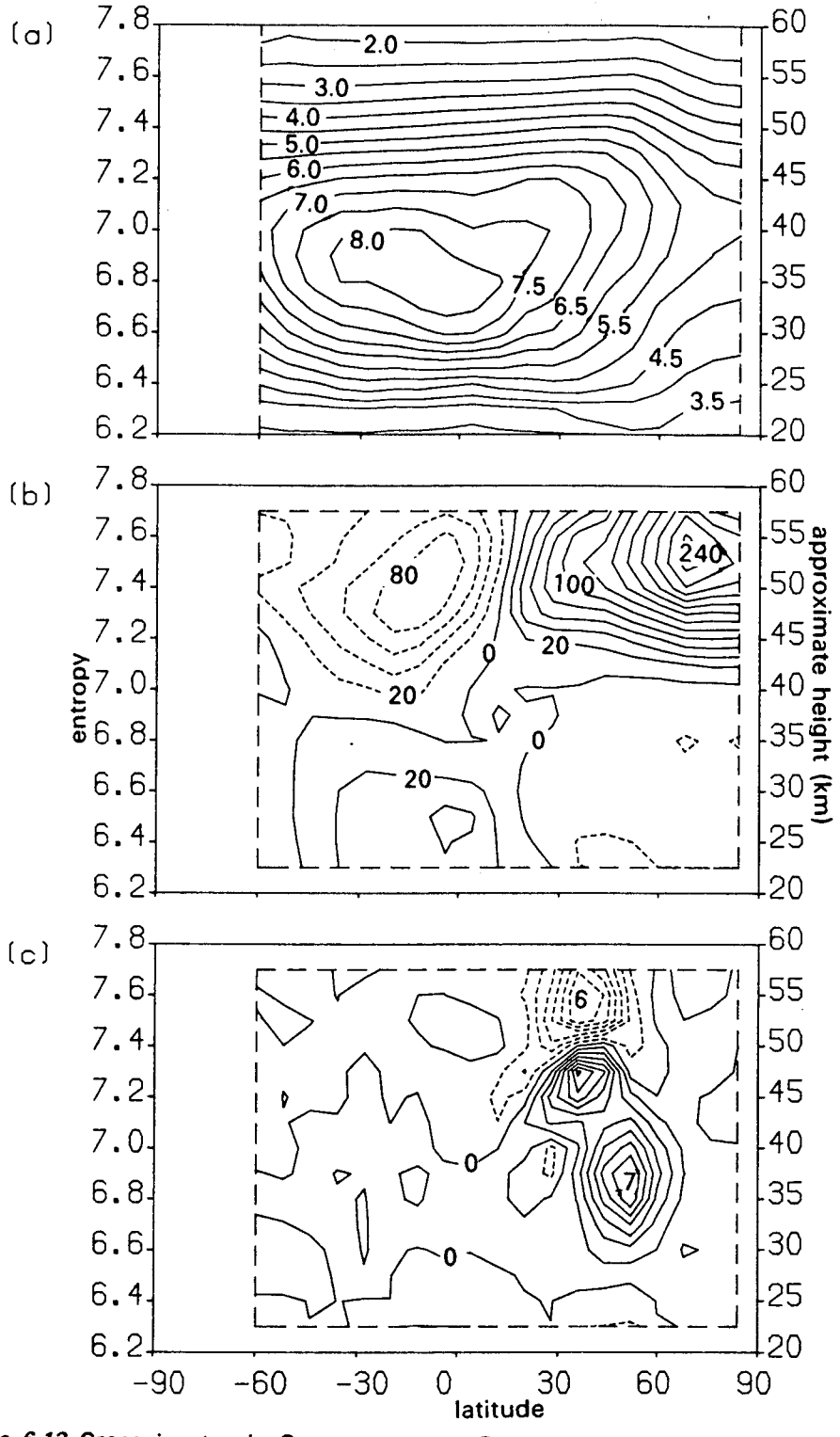


Figure 6.13. Cross-isentropic Ozone transport, December 1978. (a) the mass-weighted zonal-mean ozone distribution (ppmv, contour interval 0.5ppmv); (b) the zonal-mean cross-isentropic advection of ozone (ppbvday^{-1}, contour interval 20ppbvday^{-1}); (c) the zonal-mean eddy flux of ozone across isentropic surfaces (ppbvday^{-1}, contour interval 1ppbvday^{-1}). Again, solid lines are positive, dashed lines negative.

winter hemisphere and at high levels in the tropics and summer hemisphere but negative elsewhere (figure 6.13b). The strength of the descending motion at high levels near the winter pole is such that the advection is almost 250ppbv day^{-1} near 75°N at η=7.6; in this region the mixing ratio is around 2.5ppmv, so we see that this is decreased at a rate of 10% per day by the vertical transport, this change must be offset by other mechanisms. The maximum negative advection is 90ppbvday^{-1} near 10°S, η=7.4.

By comparison, the cross-isentropic eddy transport (figure 6.13c) is generally much weaker. It peaks at 7ppbvday^{-1} near 55°N, η=6.8, with a decrease of -7ppbvday^{-1} near 40°N, η=7.6; these changes are less than 10% of those brought about by the mean meridional advection. Indeed, at high levels, the mean advection is two orders of magnitude stronger than the eddy transport. However, near the level of ozone maximum, where the vertical gradient of the mean field changes sign, the mean transport is very weak and the eddy transport is of comparable magnitude. Further, the two transports are of opposite sign at this level, so it seems that there could be a near-cancellation between the mean and eddy transport below η=7.0.

The corresponding fields for February (figure 6.14) show similar features. The vertical mean advection (figure 6.14b) is somewhat weaker than in December as the descent over the pole is weaker (as discussed in section 6.4, figures 6.7 and 6.9), although the pattern of zonal-mean ozone mixing ratio (figure 6.14a) is not substantially different from that of December (figure 6.13a). The strongest downward eddy transport of 10ppbvday^{-1} at 60°N, η=7.6 (figure 6.14c) is around 20% of the magnitude of the mean upward transport there. Again at lower levels in the winter stratosphere the eddy transport is of comparable magnitude to the mean advection.

Results for other months are similar to these: the mean vertical advection is at least one order of magnitude stronger than the eddy transport at high levels, whilst in the lower middle stratosphere the mean advection is around five times as strong as the eddy transport of ozone, the two terms being comparable near the level of ozone maximum in the middle stratosphere.

One interesting aspect of these results is the position of the maximum in vertical eddy transport. In November 1978 (not shown) the vertical eddy transport was 4ppbvday^{-1} near 35°N above η=7.2; this strengthened in

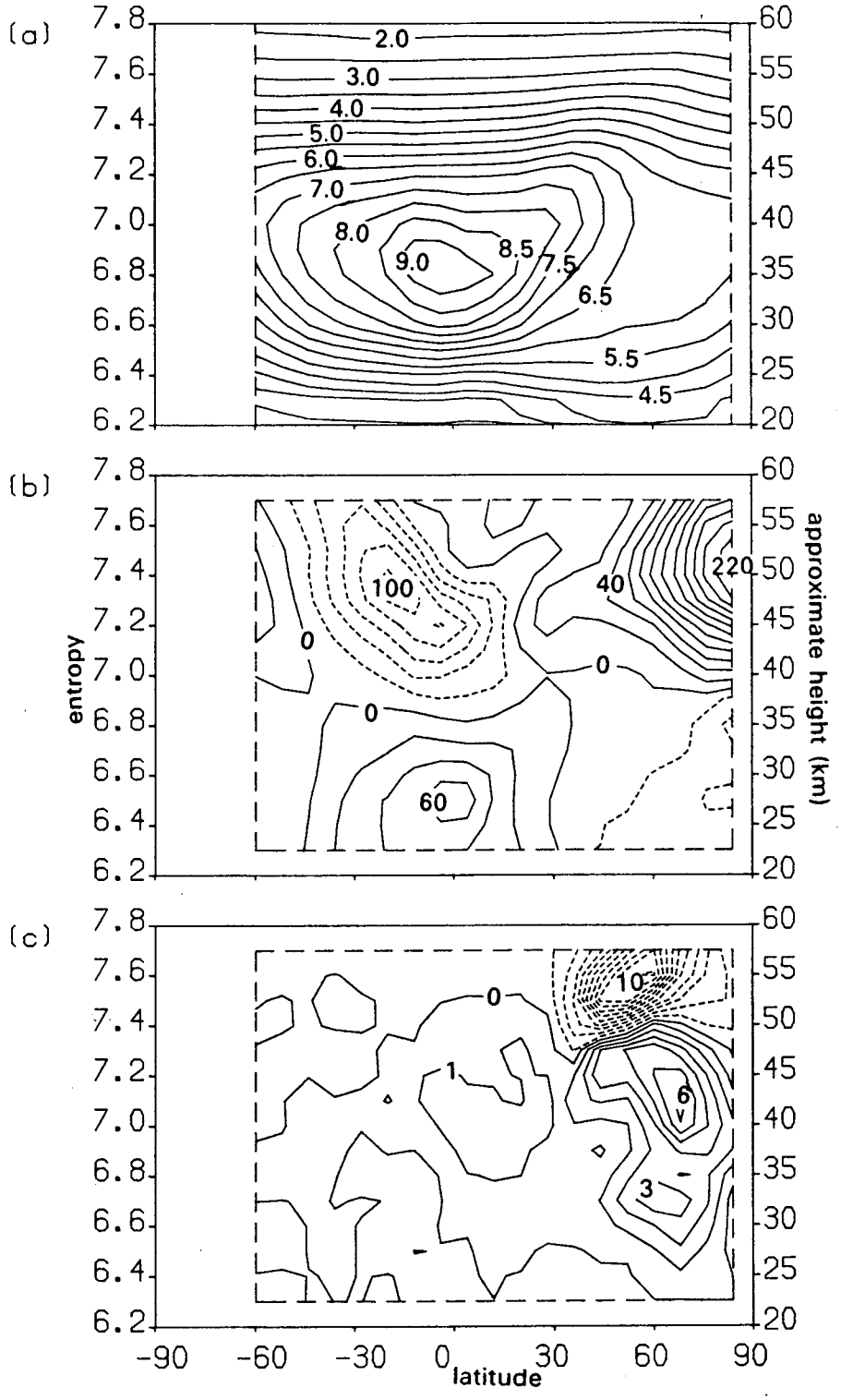


Figure 6.14. Cross-Isentropic Ozone transport, February 1979. Details as in figure 6.13.

December (figure 6.13c), moving only slightly poleward but penetrating lower into the stratosphere. In January 1979 (not shown) the vertical eddy transport (6ppbvday^{-1}) was in a deep layer between η=6.6 and 7.4 which sloped polewards with increasing height. By February (figure 6.14c) the eddy transport was even closer to the pole and was still over a deep layer. In March, when the wintertime stratospheric structure finally broke down, it was much weaker and confined to higher levels; it was negligible in the northern hemisphere in April and May. This closing in of the eddy-transport to the pole through the winter is a consequence of the strongest zonal-asymmetries in the radiative heating distribution being concentrated at the edge of the polar vortex; as it eroded away through the winter by the nonlinear dynamics of Rossby wave breaking, and moved away from the pole, the eddy-activity moved closer to the pole. This observation is consistent with the expectation that the strongest non-zonalities in the radiative heating field are correlated with the large deviations from zonal-asymmetry of the ozone distribution.

6.5.4. Mean and Eddy Water Vapour Transport.

Before drawing definite conclusions about the role of cross-isentropic eddy transport in the atmosphere we should calculate this term for a tracer with an atmospheric distribution different to that of ozone. Tracers such as methane and nitrous oxide have tropospheric sources and their mixing ratio decreases with altitude in the stratosphere; water vapour mixing ratio tends to increase with height because of its mesospheric source. Because LIMS data includes three-dimensional fields of water vapour, this has been used for this test.

The mass-weighted zonal-mean distribution of water vapour (figure 6.15) shows how it increases with height in the stratosphere. Its mean, cross-isentropic advection (figure 6.15b) is consistent with downward transport in the polar region of the northern hemisphere and upward transport in the tropics. The eddy flux (figure 6.15c) is generally around an order of magnitude smaller than the mean advection. This reaffirms the notion that the diabatic eddy transport may be neglected in relation to the cross-isentropic transport by the zonal-mean diabatic circulation.

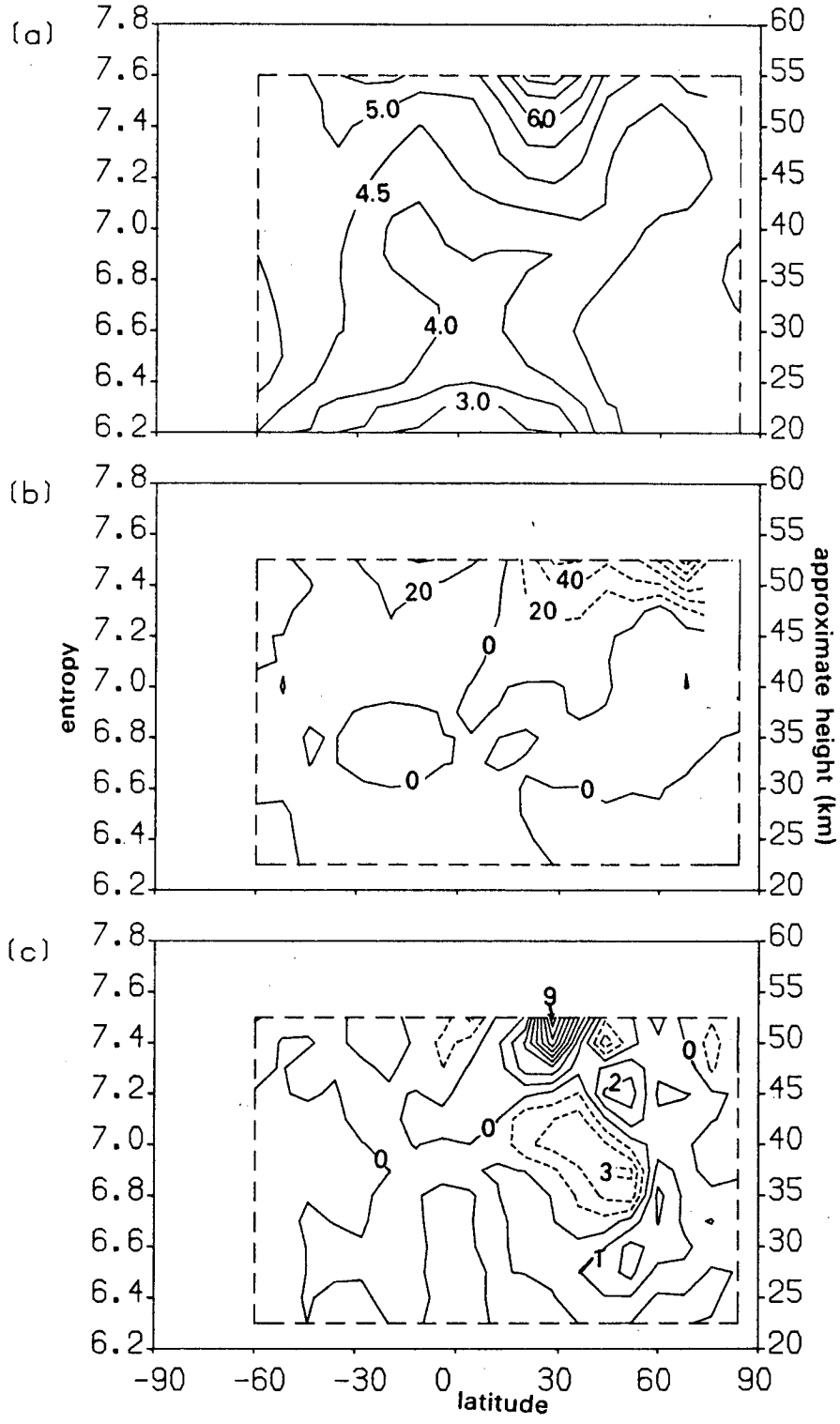


Figure 6.15. Cross-Isentropic Water vapour transport, December 1978. Details as in figure 6.13.

6.5.5. Conclusions.

The results of this section have shown that for two tracers with different vertical profiles the mean cross-isentropic transport generally dominates the eddy contribution by more than an order of magnitude, and often two orders of magnitude; the only exception occurred for ozone in the region of the middle stratosphere where its vertical gradient changes sign, so that its mean advection almost disappears. The conclusion to be drawn from this is that the diabatic eddy transport may (usually), to a very good approximation, be neglected in comparison to the mean diabatic transport. Thus, any systematic cross-isentropic transport of trace gases across isentropic levels is by the diabatic circulation.

6.6. Implications For Zonal-Mean Modelling Studies.

The results of the last section, along with those of Tung (1982) who showed that the tracer transport by planetary waves is predominantly along isentropic surfaces, imply that the diabatic eddy transport term in the zonal-mean tracer budget equation may be neglected. This is true only in isentropic coordinates, since in pressure or height coordinates an adiabatic eddy causes undulations in the height of isentropic levels which leads to a vertical eddy transport if the perturbation is not reversible.

Consider now the K-theory developed in section 6.2. Since $\partial \Lambda_\eta / \partial \eta \simeq 0$, equations (6.10), (6.11) and (6.13), along with the knowledge that neither the meridional nor the vertical gradient of the zonal-mean tracer mixing ratio vanishes identically, we must have $K_{\eta\eta}=D_{\eta\eta}\equiv 0$ and $D_{\eta\phi}=L_{\eta\phi}\equiv 0$. By the symmetric and anti-symmetric nature of D and L we must then have $D_{\phi\eta}=L_{\phi\eta}\equiv 0$. Then, the only non-vanishing eddy transport coefficient is the isentropic diffusion term, $D_{\phi\phi}$. Since $L_{\phi\eta}\equiv 0$ the eddy advection velocity defined by equation (6.14) must vanish. Thus, the zonal-mean tracer transport equation (6.15) may be simplified to:

$$\frac{\partial}{\partial t}(<\mu>) + \frac{<v>}{a}\frac{\partial}{\partial \phi}(<\mu>) + <w>\frac{\partial}{\partial \eta}(<\mu>) = <S>$$

$$+([\sigma_\eta])^{-1}\left(S_E - \frac{1}{a\cos\phi}\frac{\partial}{\partial \phi}\{[\sigma_\eta]K_{\phi\phi}\frac{\partial}{\partial \phi}(<\mu>)\right). \qquad (6.16)$$

Thus, only one eddy transport coefficient is required, making the problem much more straightforward than in the unapproximated case. If this diffusion coefficient can be calculated the problem will be soluble.

Holton (1986) and Tung (1986) have shown that $K_{\phi\phi}$ can be obtained from a knowledge of the dynamics of the system: it is directly proportional to the divergence of the Eliassen-Palm flux or the northward eddy flux of potential vorticity. A fully consistent zonal-mean model would have a diffusion coefficient determined from the evolving zonal-mean model state; however, a method for calculating $D_{\phi\phi}$ internally has not yet been accomplished, so most current models use specified transport tensors. This could affect the results obtained with such models, since the propagation of planetary waves depends upon the state of the atmosphere and could change significantly if the zonal-mean atmosphere were perturbed from its current state.

The importance of the eddy diffusion was considered by Holton (1986), in a study of the transport of trace gases, such as methane and nitrous oxide, in the lower stratosphere. He assumed the relationship justified here: that the eddy transport was isentropic so that equation (6.16) exactly describes the flow. If advection were simply by a diabatic circulation the distribution of such tracers (with tropospheric sources) would be determined by the latitudinal variation of the diabatic circulation; their slope relative to isentropic surfaces would be determined by latitudinal variations in the strength of the diabatic circulation. Using a simple model, with ascent in the tropics and descent at the pole Holton showed that the slope of the isopleths of tracer mixing ratio was reduced by the action of and eddy diffusion coefficient, which tends to mix the gas horizontally; this is shown schematically in figure 6.16.

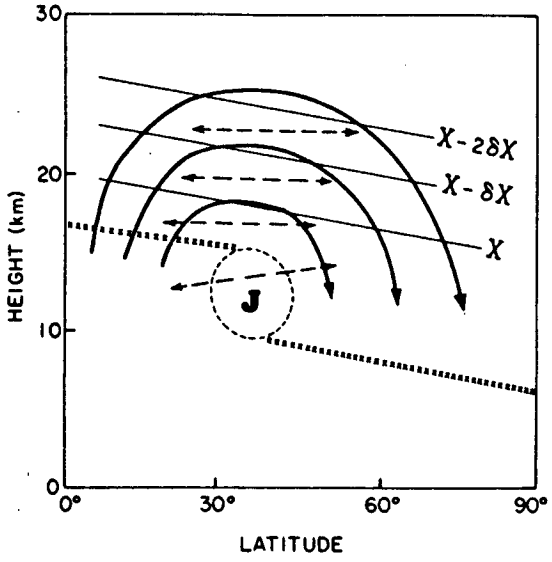


Figure 6.16. A schematic meridional cross-section illustrating transport in the stratosphere. Heavy lines with arrows show the streamlines of the diabatic circulation. Dashed arrows indicate the quasi-isentropic eddy transport. The mean tropopause is indicated by crosses and the J indicates the mean location of the subtropical tropospheric jet stream. Light lines labelled with mixing ratio values X (corresponding to μ in the main text) show the mean slope of a typical vertically stratified tracer. (From Holton, 1986.)

Holton also showed that the maximum slope for the tracer mixing ratio occurs when the timescale for the eddy diffusion lies between the radiative and photochemical timescales. This shows the importance of specifying the eddy diffusion coefficient consistently within zonal-mean model of the middle atmosphere, since an incorrect tracer distribution would result if its timescale did not bear the correct relationship to the radiative and photochemical timescales. As yet, no isentropic coordinate, zonal-mean model has been constructed with an internally calculated diffusion coefficient. The isentropic coordinate model of Ko *et al.* uses a specified $K_{\phi\phi}$. This should not be too important for short-term integrations, but when the zonal-mean model is applied to climatological problems there is no guarantee that after, say, several decades the eddy forcing of the atmosphere may not change in nature. Fully consistent models should calculate their own diffusion coefficients.

There are further practical problems concerned with using zonal-mean

models in isentropic coordinates. Radiative and photochemical processes are pressure-dependent, so that they need to be performed on grids with known pressures. The most efficient manner of doing this would seem to be continual interpolation between isentropic coordinates for the dynamics and pressure coordinates for the physics, which could require lots of computational storage space and resources. Practical use of isentropic coordinate models may be restricted by these problems.

CHAPTER 7

DIAGNOSES OF THE THREE-DIMENSIONAL DIVERGENT MOTION.

7.1. The Dominance of the Diabatic Component of Divergence.

In Chapter 2 it was shown that the divergent horizontal mass flux may be partitioned into diabatic and transient components, respectively associated with vertical gradients of transport across isentropic surfaces and the transient rearrangement of mass within isentropic layers (equations (2.41) and (2.42)). In the zonal-mean the diabatic component is dominant over long periods, so that the monthly-averaged mean-meridional-circulation is well approximated by the diabatic circulation.

It is not immediately clear that this same balance of terms will apply in the three-dimensional case. To illustrate this consider the situation of a winter hemisphere in which the monthly-mean vortex is displaced from the pole to a different location between two months (e.g. January and February 1979 at η=6.7, figures 5.16a and 5.17a). The change in position of the vortex causes a rearrangement of the mass within the isentropic layer as well as changes in the height of the level; if these do not compensate for each other, $\partial\sigma_\eta/\partial t$ could be important; such redistribution of matter would not leave a signal in the zonal-mean. The important question is whether this rearrangement of matter leads to a transient forcing with magnitude comparable to the diabatic forcing when averaged over the month.

The two forcing terms have been examined using LIMS data. For January 1979 at η=6.7 (figure 7.1) the diabatic forcing is generally an order of magnitude greater than the transient forcing (which was calculated from the difference between the distributions of density in February and December); this dominance holds throughout the northern winter of 1978-79. Thus, to a good approximation we deduce that the diabatic divergent motion is dominant in the monthly-mean. This chapter is devoted to this three-dimensional diabatic circulation; the method of solution and the accuracy of the derived circulation are considered before the results are discussed.

For the purpose of analysing the results it is convenient to introduce a quasi-divergent velocity, $\mathbf{u}_\chi=\nabla_\eta\chi/\sigma_\eta$, for comparison with the geostrophic velocity. Note that if the density at each isentropic level were constant this

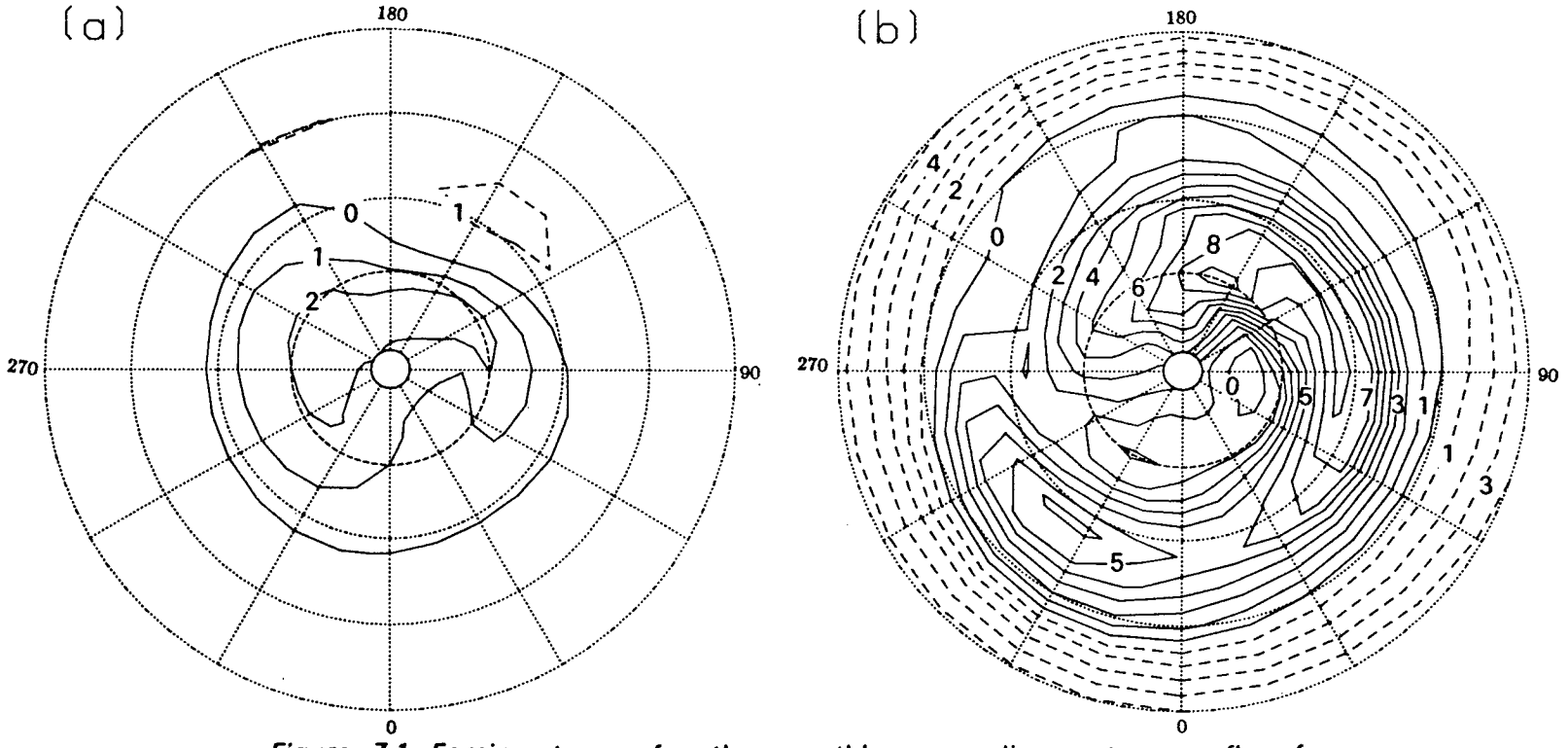


Figure 7.1. Forcing terms for the monthly-mean divergent mass-flux for January 1979 at η=6.7. (a) $\partial\sigma_\eta/\partial t$ and (b) $\partial W/\partial\eta$; both have units $kgm^{-2}day^{-1}$ and contour interval $1kgm^{-2}day^{-1}$ and are shown between 4°N and 84°N.

velocity would be exactly the divergent velocity; however, σ_η generally has a gradient (particularly in the northward direction) so that $\nabla_\eta\chi$ contains a component related to this and hence $\mathbf{u}_\chi$ is not exactly irrotational.

7.2. Solution of the Divergence Equation on the Sphere.

To obtain the distribution of the potential function, χ, it was necessary to invert the Laplacian operator in equation (2.41), using the known distribution of $\partial W/\partial\eta$. This was achieved by relaxation, using a routine provided by Dr. R.S. Harwood.

Since the problem was posed in spherical geometry and the extrapolated data set was global no boundary condition was needed. The χ-field was determined to within an arbitrary constant, whose value is immaterial since it did not affect $\nabla_\eta\chi$.

Given the errors in the data, an error of 1% in the velocity $\mathbf{u}_\chi=\nabla_\eta\chi/\sigma_\eta$ was deemed acceptable. To obtain this accuracy in χ it was found necessary to obtain χ accurate to 8 significant figures; this was determined by calculating χ correct to 13 significant figures and determining $\mathbf{u}_\chi$ from this field; the χ-solution was then successively degraded until $\mathbf{u}_\chi$ changed by 1% at any gridpoint.

7.3. The Accuracy of the Calculations.

The accuracy of the potential function χ_D depends upon the error in the vertical gradient of the vertical mass-flux rather than on that in the flux itself. Since the flux is determined from the net radiative heating rate, which tends to have a 'smooth' vertical structure because of the dependence of the irradiance at any level upon the atmospheric state at other levels, its vertical gradient is much less sensitive to errors in the data than the flux itself. Consequently, an error of 20% seems to be a reasonable upper limit for the error in $\partial W/\partial\eta$, even though the heating rate has an error of ±25%.

Solution of equation (2.41) involves inversion of a Laplacian operator, as already discussed. In order to determine how an error in $\partial W/\partial\eta$ may manifest itself into χ several assumptions seem necessary. Writing $R=\partial W/\partial\eta$, assume that it can be expressed in terms of a Fourier series:

$$R = \sum_{m,n}(\Xi_{m,n} + \varepsilon_{m,n})\exp[-i(mx+ny)]$$

where Ξ represents the amplitude of the real wave and ε represents the amplitude of the error; m and n are assumed to be zonal and meridional wavenumbers for the structure, which is valid locally (but, strictly, spherical analysis should be performed with the meridional structure represented by Legendre polynomials). Since equation (2.41) is linear each wave can be treated individually. Integration then gives:

$$\chi_{m,n} + \delta_{m,n} = -(m^2+n^2)^{-1}(\Xi_{m,n} + \varepsilon_{m,n})\exp[-i(mx+ny)] \qquad (7.1)$$

(for each pair of wavenumbers). Here, χ is assumed to be the true solution (so that $\nabla_\eta\chi = R$) and δ is assumed to be the error in χ resulting from the error in the forcing term (so $\nabla_\eta\delta = \varepsilon$). Consequently, the fractional error in the solution for each wavenumber is:

$$\delta/\chi = \nabla_\eta^2\varepsilon/\nabla_\eta^2\Xi$$
$$= \varepsilon/\Xi$$

which is simply the error in the forcing term.

Thus, if $\partial W/\partial\eta$ is known to within 20% at each wavenumber, so is χ.

Thus, the analysis so far has shown that the error at each wavenumber is not increased by the integration. Of more profound importance for our purposes is the accuracy of the entire field, obtained by the summation of equation (7.1) over all wavenumbers; the contribution of each wave to the entire field has an inverse-dependence on wavenumber, so that the shorter waves have less signal in the integrated field than in R (this is the well-known relationship that inverting a Laplacian smooths the field). Thus, the distribution of χ is dominated by the lower wavenumbers. If noise is assumed to be random it will have a stronger signal at the higher wavenumbers in a Fourier analysed field, so that integration of this field will suppress the noise.

Consequently, the error in χ is not worse than the error in $\partial W/\partial\eta$.

7.4. The Divergent Flow at η=7.5

The eastward geostrophic flow at this level showed a strong jet near 40°N from November until January, but the flow was weaker in February and March. As an example, the zonal flow in December (figure 7.2) exceeded 100ms^{-1} between 330°E and 60°E at 40°N; it was considerably weaker than this in the opposite part of the northern hemisphere. The flow in November was rather weaker than that in December because the strong meridional temperature gradients had not formed. In January it was of comparable strength to December but it was shifted slightly polewards following the Canadian warming (Labitzke, 1981).

The zonal quasi-divergent velocity, u_χ, was considerably weaker than the zonal geostrophic flow. The eastward component did not exceed a few tenths of 1ms^{-1}, which is around two orders of magnitude smaller than the zonal geostrophic flow; it thus represented a very small correction to the eastward velocity field. By contrast, the northward component of the quasi-divergent flow, v_χ, could exceed 10% of the magnitude of the northward geostrophic velocity; the northward components of the flow are now discussed in more detail.

a. November 1979, figure 7.3 The strongest regions of northward flow were 10ms^{-1} near 330°E, 70°N and -15ms^{-1} near 240°E,70°N. The maximum in v_χ (0.8ms^{-1}) occurred in a broad region between 150-240°E, 46-70°N; the geostrophic flow in this region was typically 5-10ms^{-1}, so that v_χ was around 10-20% of v_G.

b. December 1978, figure 7.4. The wavenumber-1 structure to the northward motion, with flow over the polar cap (20ms^{-1} near 90°E and -15ms^{-1} near 270°E is in accord with the potential vorticity distribution (figure 5.7a). The strongest quasi-divergent northward flow was located at 180°E, 55°N, its magnitude of 1.6ms^{-1} being more than 10% as strong as the geostrophic flow (10ms^{-1}) at that location. In the region of strongest northward geostrophic flow the divergent component almost vanished. It is of considerable interest that this happens: there is a distinct anti-correlation between the distributions of v_G and v_χ near the North pole, with maxima in v_χ occurring in regions of minimum v_G, and *vice-versa*.

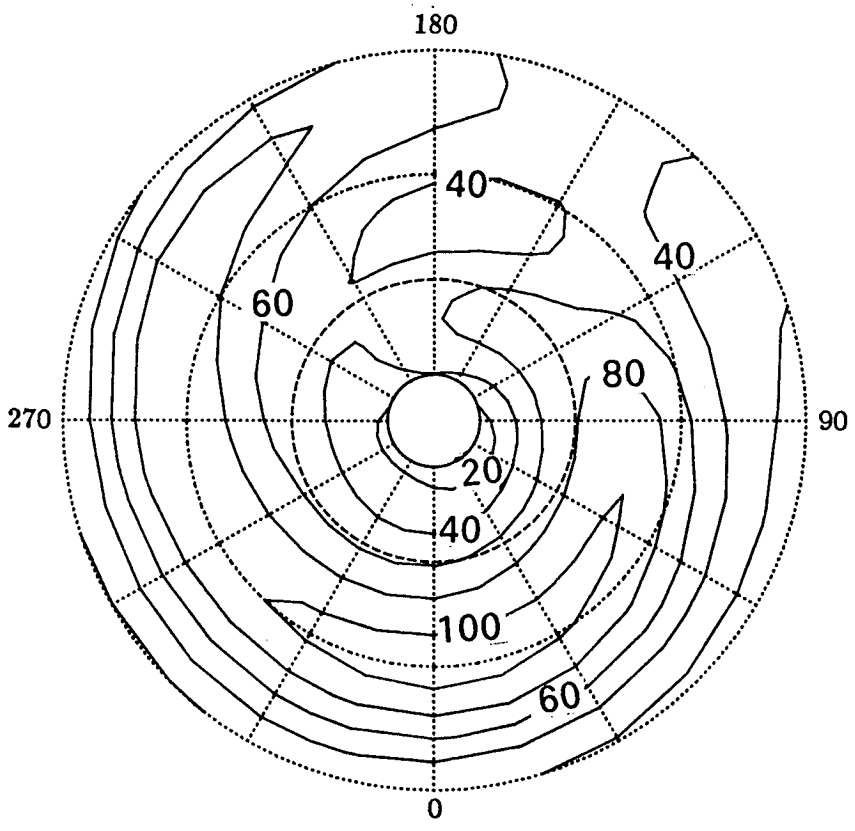


Figure 7.2. The monthly-mean zonal (eastward) geostrophic velocity (ms^{-1}) at η=7.5 for December 1978 between 20°N and 80°N. The contour interval is $20ms^{-1}$.

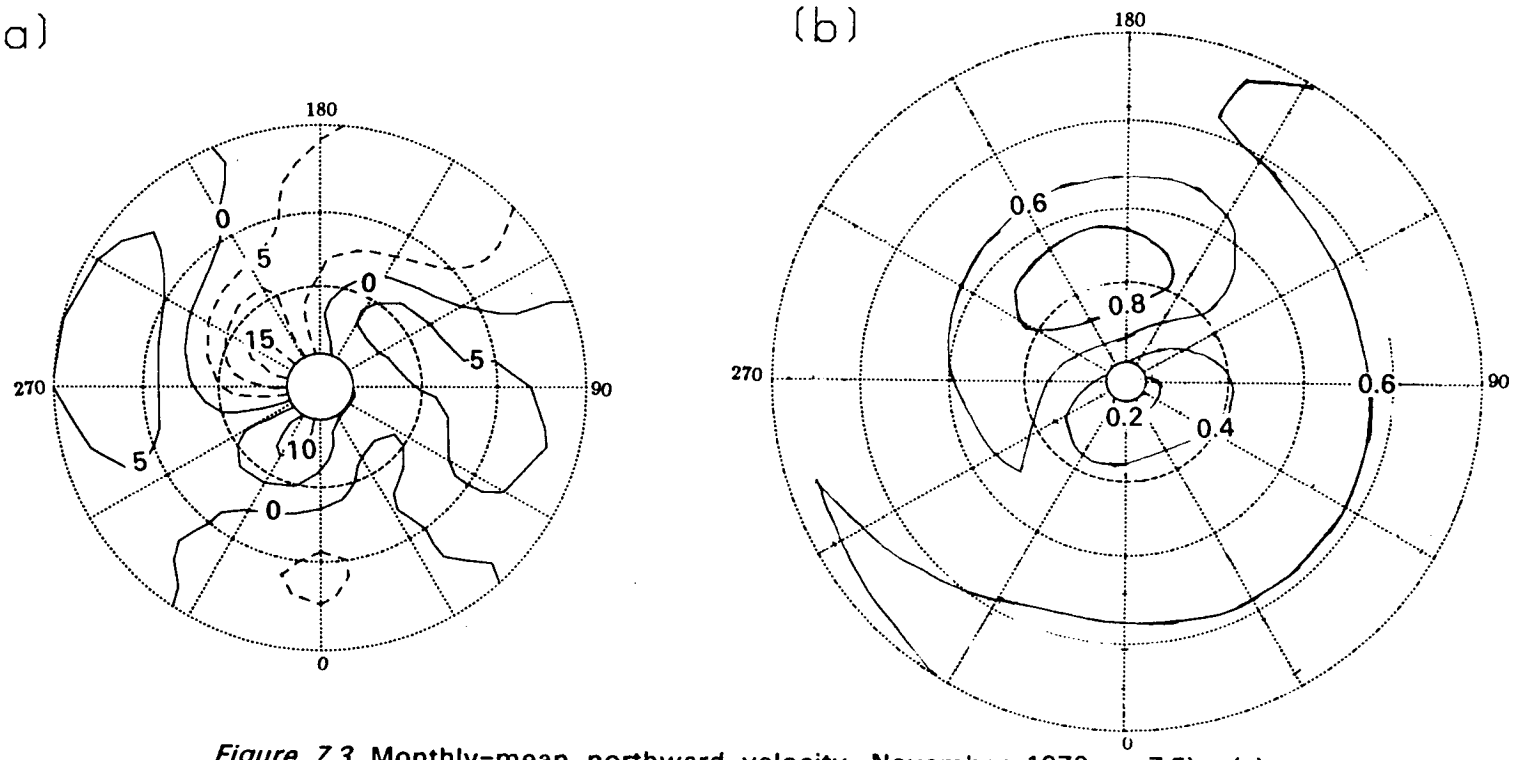


Figure 7.3. Monthly-mean northward velocity, November 1978, η=7.5). (a) v_G (ms^{-1}, contour interval $5ms^{-1}$) between 20°N and 80°N; (b)v_χ (ms^{-1}, contour interval $0.2ms^{-1}$) between 4°N and 84°N.

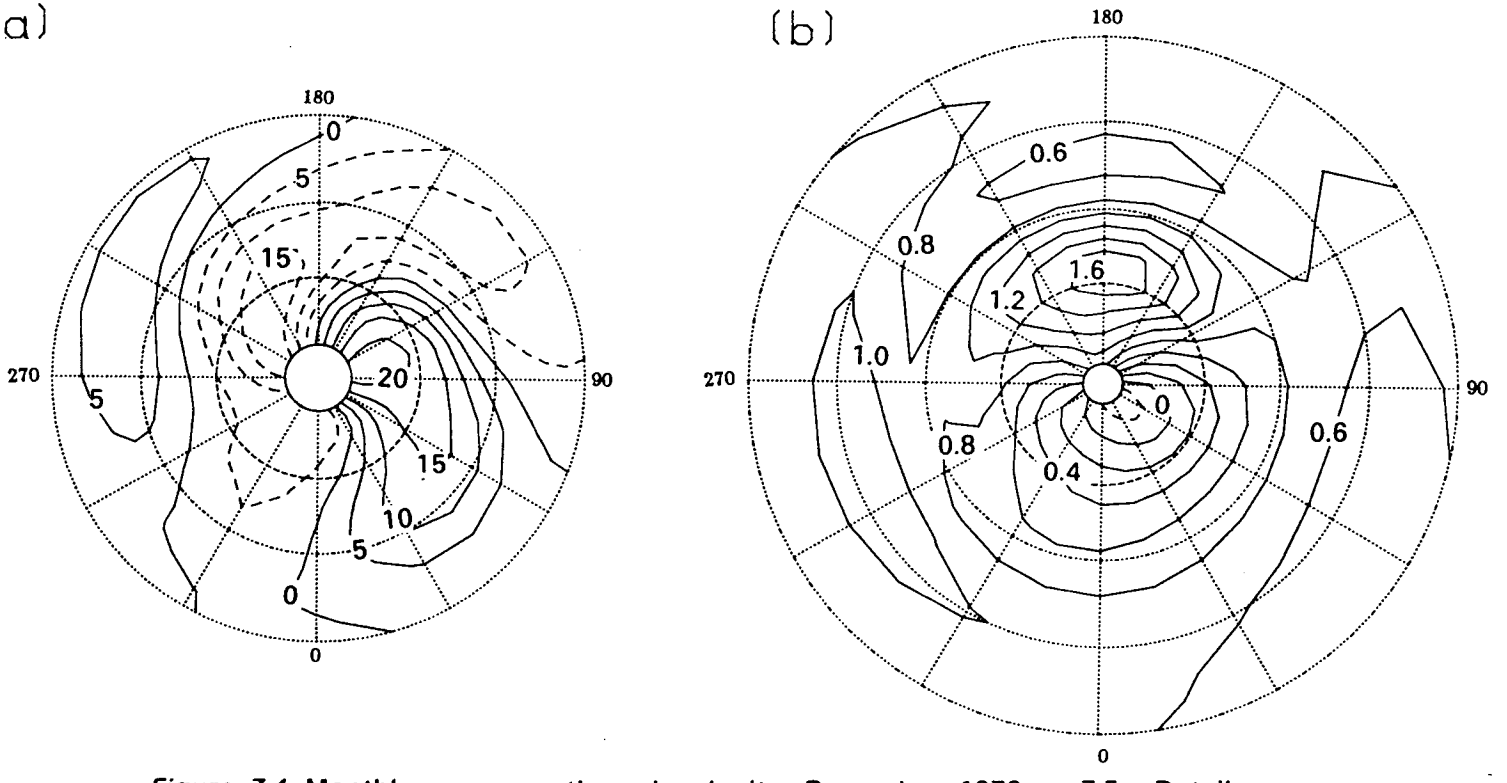


Figure 7.4. Monthly-mean northward velocity, December 1978, η=7.5. Details as for figure 7.3.

c. January 1979, figure 7.5. The monthly-mean polar vortex was more distorted than in December (compare figures 5.7a and 5.9a); there was strong northward flow to each side of the vortex, particularly on its eastern flank. A maximum v_χ of 1.8 ms^{-1} occurred near 100–180°E, 40°N, where v_G ranged between 0–10ms^{-1}; again it is evident that v_χ was of the order 10–20% of v_G. It is of interest to note that this occurred in the region of low potential vorticity, where the Rossby number is small.

d. February 1979, figure 7.6. A strong jet was evident in the northerly flow; this reached 50ms^{-1} at 60°E,80°N on the edge of the displaced polar vortex (figure 5.10a). In this region v_χ was small (0.4ms^{-1}), but again it showed a maximum (1.8ms^{-1}) where v_G was small (<5ms^{-1}) near 160–220°E,60°N in the region of low potential vorticity. A second maximum in v_χ (1.6ms^{-1}) occurred near 330°E, 70°N, but here $v_G \simeq 20ms^{-1}$.

e. March 1979, figure 7.7. The high stratosphere was much less disturbed this month, the strongest v_G being 15ms^{-1}, with flow across the pole from 0–80°E; the divergent flow was small in this region, but reached 0.8ms^{-1} at 180–240°E, 75°N where v_G ranged from 0–10ms^{-1}.

f. April and May 1979, not shown. In these early summer months v_χ showed very little departure from zonality; it was weak and directed away from 40°N. The northward geostrophic flow showed some longitudinal structure in April but was almost zonal in May; v_χ was typically 10–20% of the magnitude of v_G.

7.5. The Divergent Flow at η=6.7

Although the eastward geostrophic velocity at this level was generally smaller than at the higher level, the eastward component of quasi-divergent velocity was still negligible by comparison. In December u_G exceeded 70ms^{-1} near 70–140°E, 50–60°N but was small at 80°N; it was of comparable magnitude in January; by February the maximum eastward flow was less than 60ms^{-1}.

As in the previous section the northward flow structure is now discussed.

a. November 1979, not shown. The northward geostrophic velocity exhibited a weak wavenumber-2 pattern which was not echoed by v_χ, which

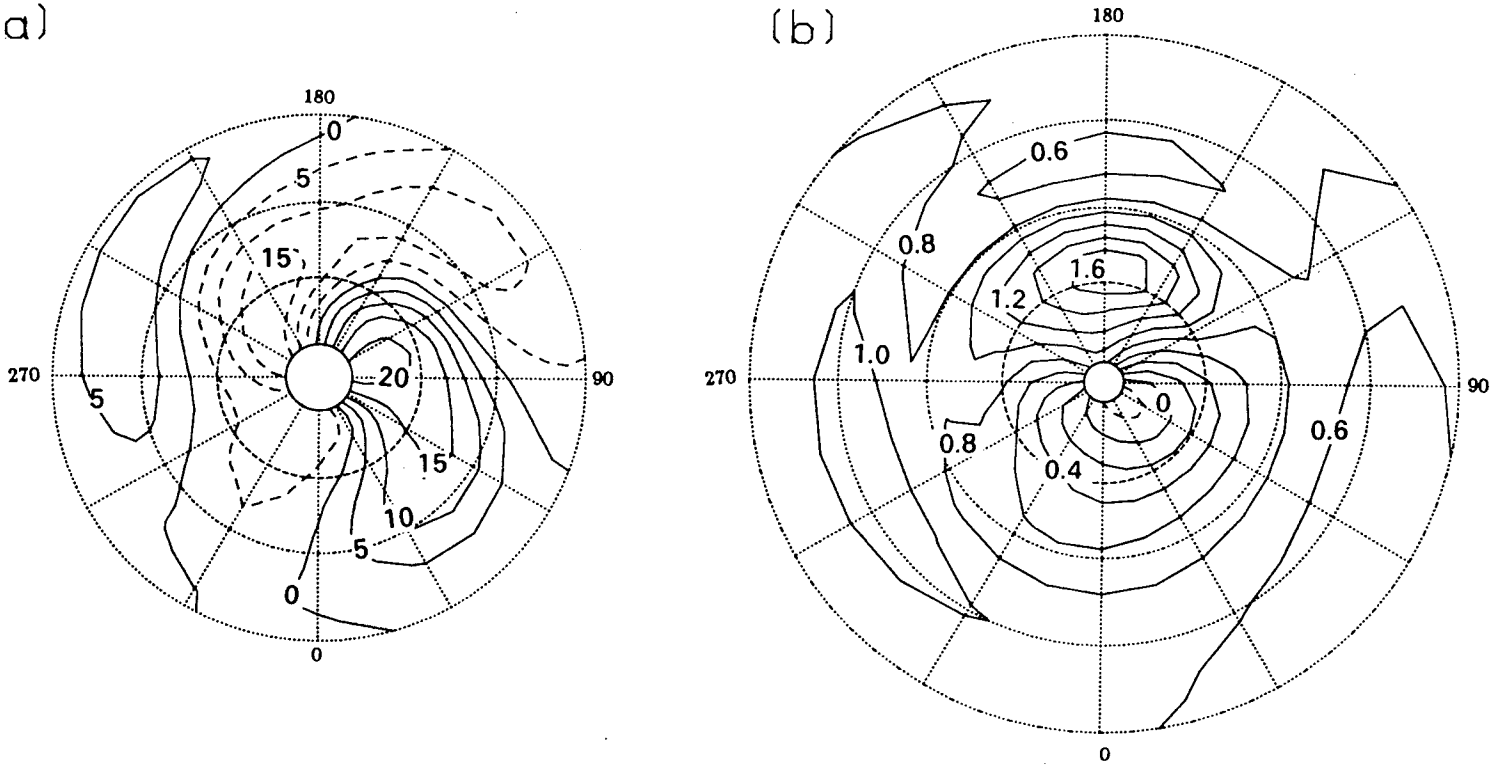


Figure 7.4. Monthly-mean northward velocity, December 1978, η=7.5. Details as for figure 7.3.

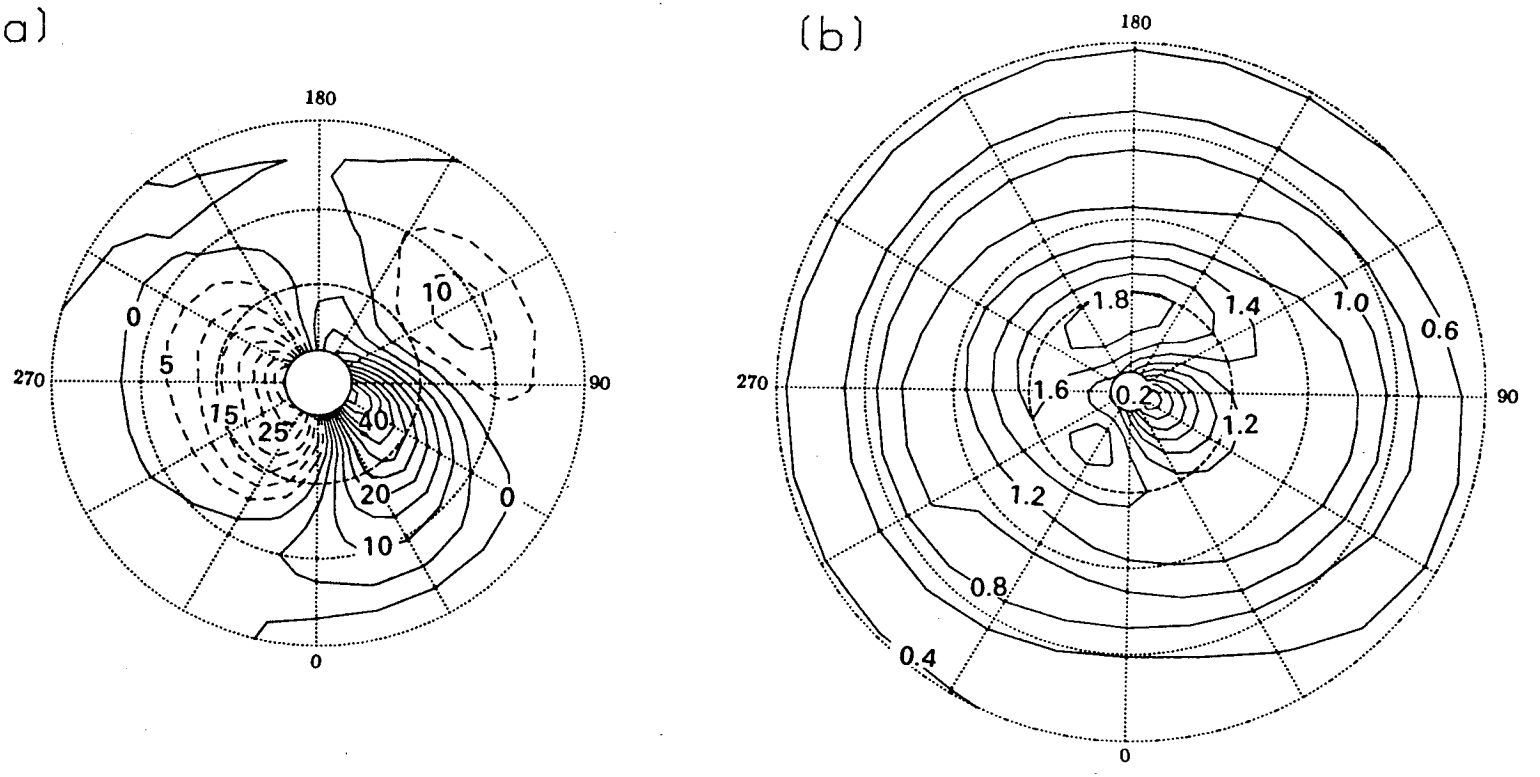


Figure 7.6. Monthly-mean northward velocity, February 1979, η=7.5. Details as for figure 7.3.

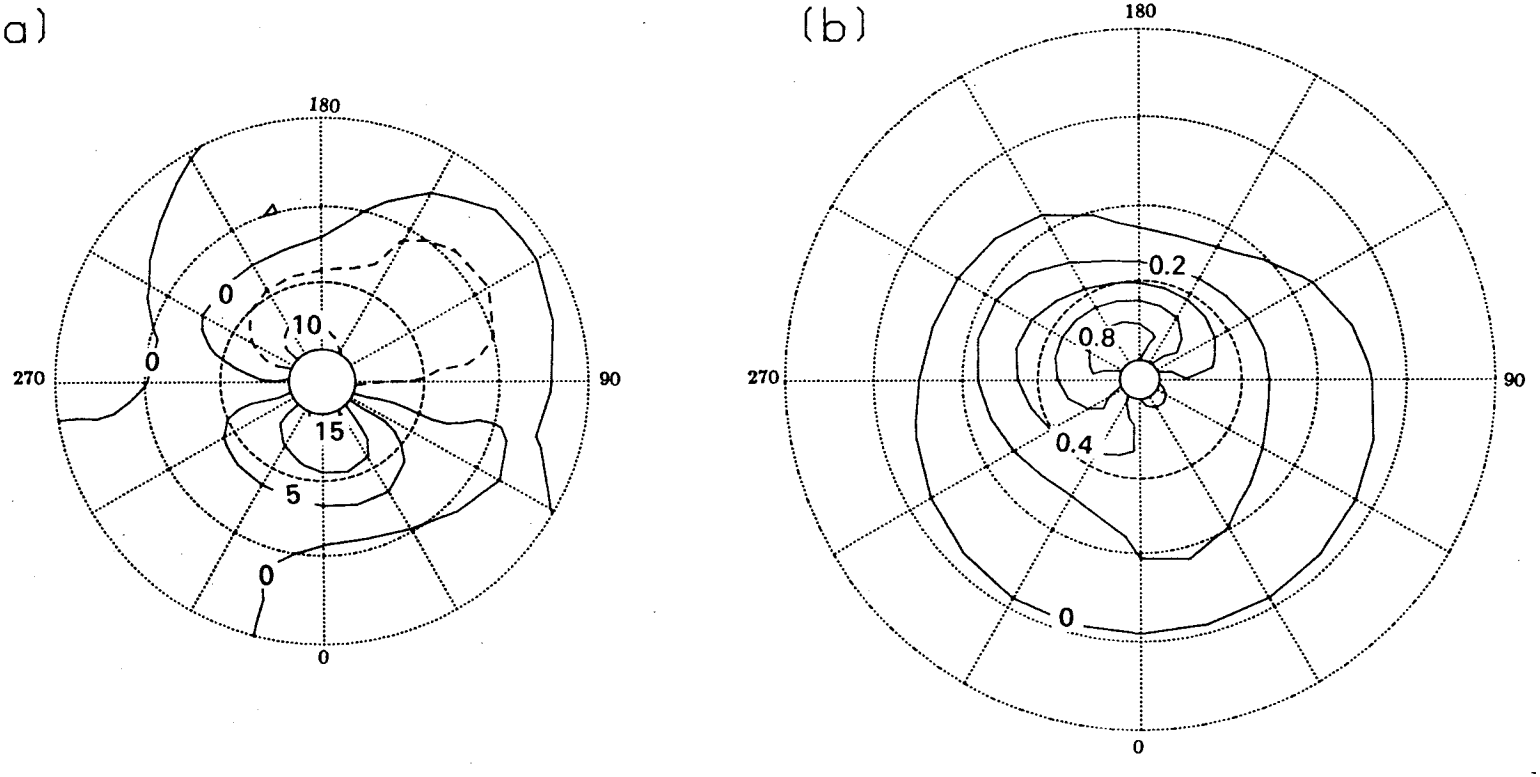


Figure 7.7. Monthly-mean northward velocity, March 1979, η=7.5. Details as for figure 7.3.

showed only weak non-zonality.

b. December 1979, figure 7.8. The maximum v_G was $20ms^{-1}$ between 120–150°E, poleward of 60°N; there was also strong southward flow ($15ms^{-1}$) near 210–240°E, 40–60°N. In each of these regions v_χ was small ($<0.3ms^{-1}$ in magnitude). The strongest v_χ ($0.5ms^{-1}$) was near 20°N over a broad range of longitudes where v_G was weaker than $5ms^{-1}$.

c. January 1979, figure 7.9. The cross-polar jet was stronger than in December, a consequence of the more highly disturbed polar vortex; v_G exceeded $40ms^{-1}$ near 160°E, 80°N. Again, v_χ was weak in the vicinity of the pole but was stronger than $0.4ms^{-1}$ between 20–40°N where v_G was weaker than $5ms^{-1}$.

d. February 1979, figure 7.10. The broad features of the previous two months were retained, although v_χ was stronger near 60°N and the increasingly displaced vortex resulted in v_G being even stronger at the pole.

e. March, April, May 1979, not shown. As the wintertime polar vortex broke down v_G weakened, becoming negligible in April and May, whilst v_χ become even more zonal in structure and turned from being northward in February to southward in May.

7.6. Discussion.

The results of this chapter have revealed that, in certain regions, the northward component of the quasi-divergent horizontal velocity, v_χ, can reach 20% of the magnitude of the geostrophic flow. This was more noticeable at the higher level (η=7.5; z≈52km) near the stratopause than in the middle stratosphere. At this level the strongest v_χ tended to occur in the region of weakest v_G; this is also the region of small Rossby number, which should imply that the geostrophic flow is a good approximation to the real wind here. The implication is that the quasi-divergent velocity is important at the higher level: there certainly appears to be a significant component of the meridional flow which cannot be detected using small-Ro theories.

The normal manner of diagnosing the divergent motion from observations is to use the vorticity budget, inferring the divergence from the residual terms in its balance (this can then be used with the continuity equation and an

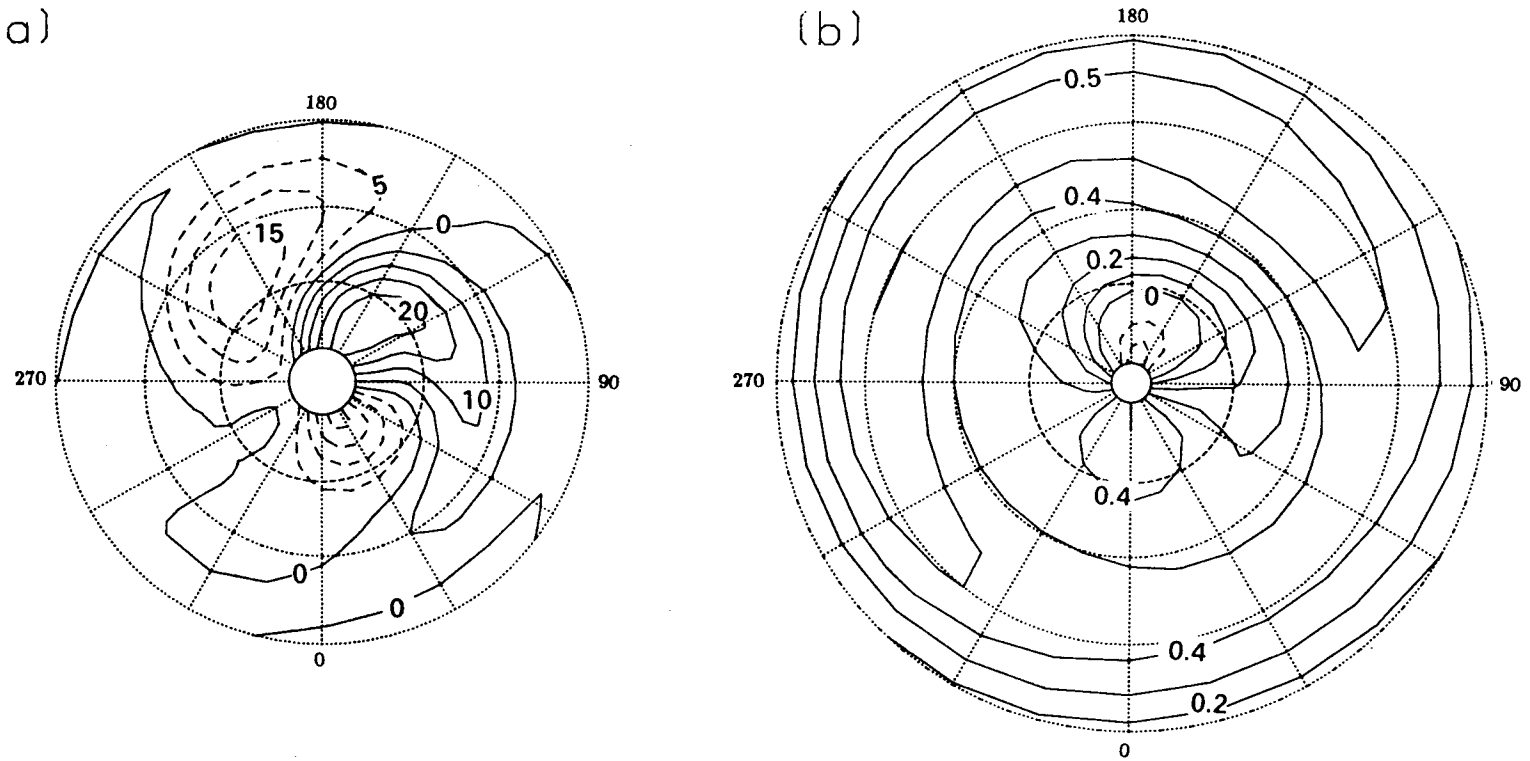


Figure 7.8. Monthly-mean northward velocity, December 1978, η=6.7. Details as for figure 7.3, except that the contour interval in part (b) is now 0.1ms^{-1}.

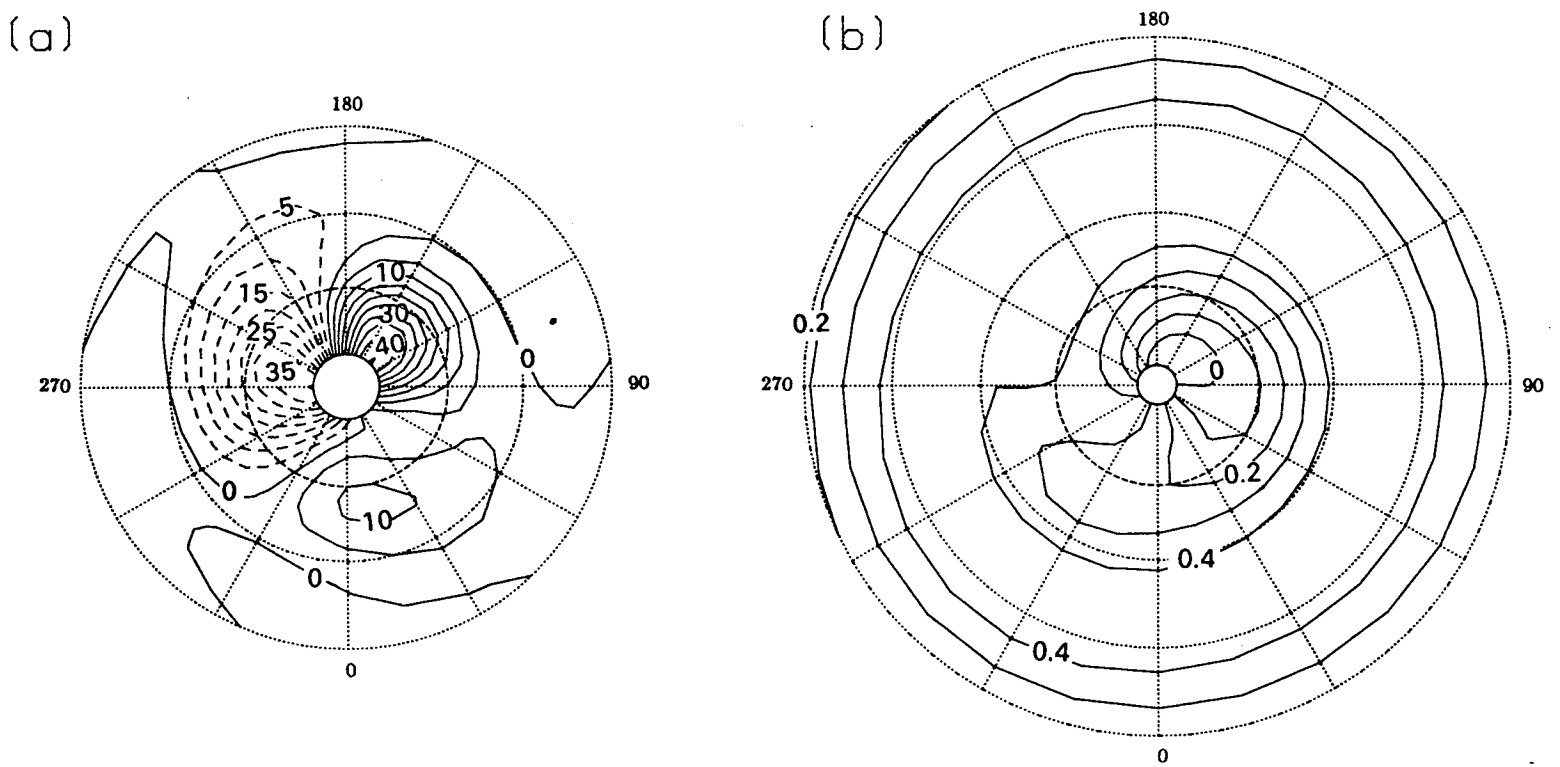


Figure 7.9. Monthly-mean northward velocity, January 1979, η=6.7. Details as for figure 7.3, except that the contour interval in part (b) is now 0.1ms^{-1}.

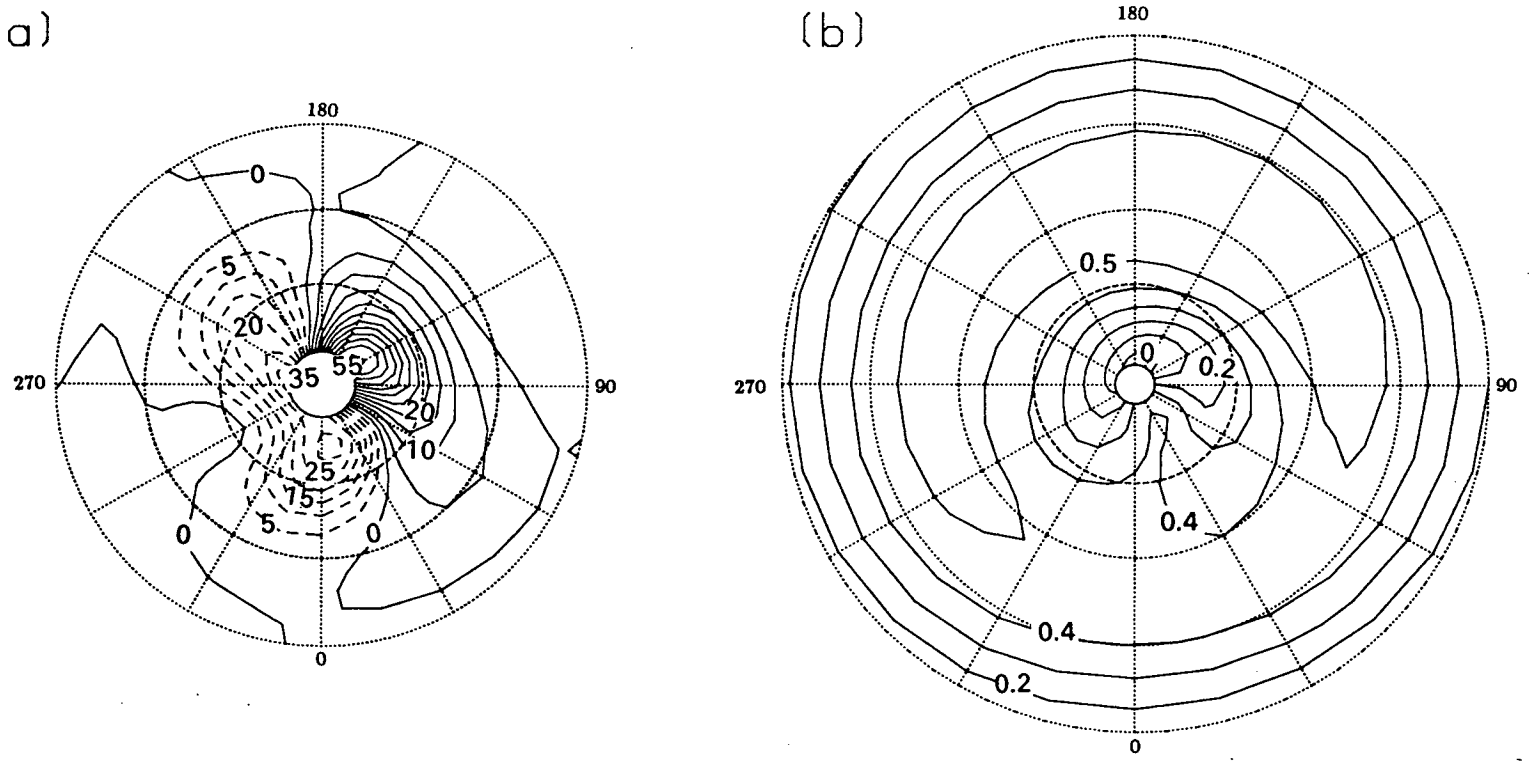


Figure 7.10. Monthly-mean northward velocity, February 1979, η=6.7. Details as for figure 7.3, except that the contour interval in part (b) is now 0.1ms^{-1}.

appropriate lower boundary condition to obtain the vertical velocity). Both diagnostic techniques are error-prone: the diabatic method used in this thesis is dependent upon the accuracy of the radiative heating rate calculations whilst the 'dynamical' method is sensitive to errors in the satellite data and the incompleteness of the rotational velocity field diagnosed from the geopotential height distribution. Because the method used in this thesis involves some integration, which tends to suppress noise, it should offer some advantages over the alternative technique, which involves far more differentiation of the temperature distributions. A comparison of the two methods to examine their consistency could be used to provide an assessment of their relative advantages and disadvantages; use of numerical model data would enable the methods themselves to be assessed in a noise-free environment, whilst their relative performances with atmospheric data would provide an estimate of their applicability to diagnostic studies.

The drawback of the 'diabatic' method of obtaining the divergent flow is that it is only valid over long time periods when the transient term of the continuity equation can be neglected. From day-to-day the transient forcing becomes important; particularly on the active days in the winter o 1978-79 the transient forcing has the same magnitude as the diabatic forcing (figure 7.11 shows the two terms for 25 January 1979). Derivation of the density, σ_η, requires differentiation and interpolation of a potentially noisy data set; despite the time-smoothing in the LIMS MAT data it is possible that determination of $\partial\sigma_\eta/\partial t$ is too erroneous for successful application, particularly with data sets other than LIMS (which tend to have a poorer vertical resolution). Thus, the technique discussed here may not be practical for calculating the divergent flow field over short periods (i.e. the daily flow).

For this reason, and because of the difficulties inherent in calculating advection terms from low temporal resolution data, a complete study of the vorticity budget has not been undertaken. At η=6.7 the divergent component of the flow seems small, so that tentatively (and qualitatively) we deduce that the flow is dominated by a rotational component; this is in accord with the work of Sawyer (1964), who found the divergence to be small near 10mbar in his study of the circulation in a quasi-steady wavenumber-1 flow in the low and middle stratosphere. At higher levels the strength of the divergent flow implies that it may play a greater role in the vorticity budget; near the

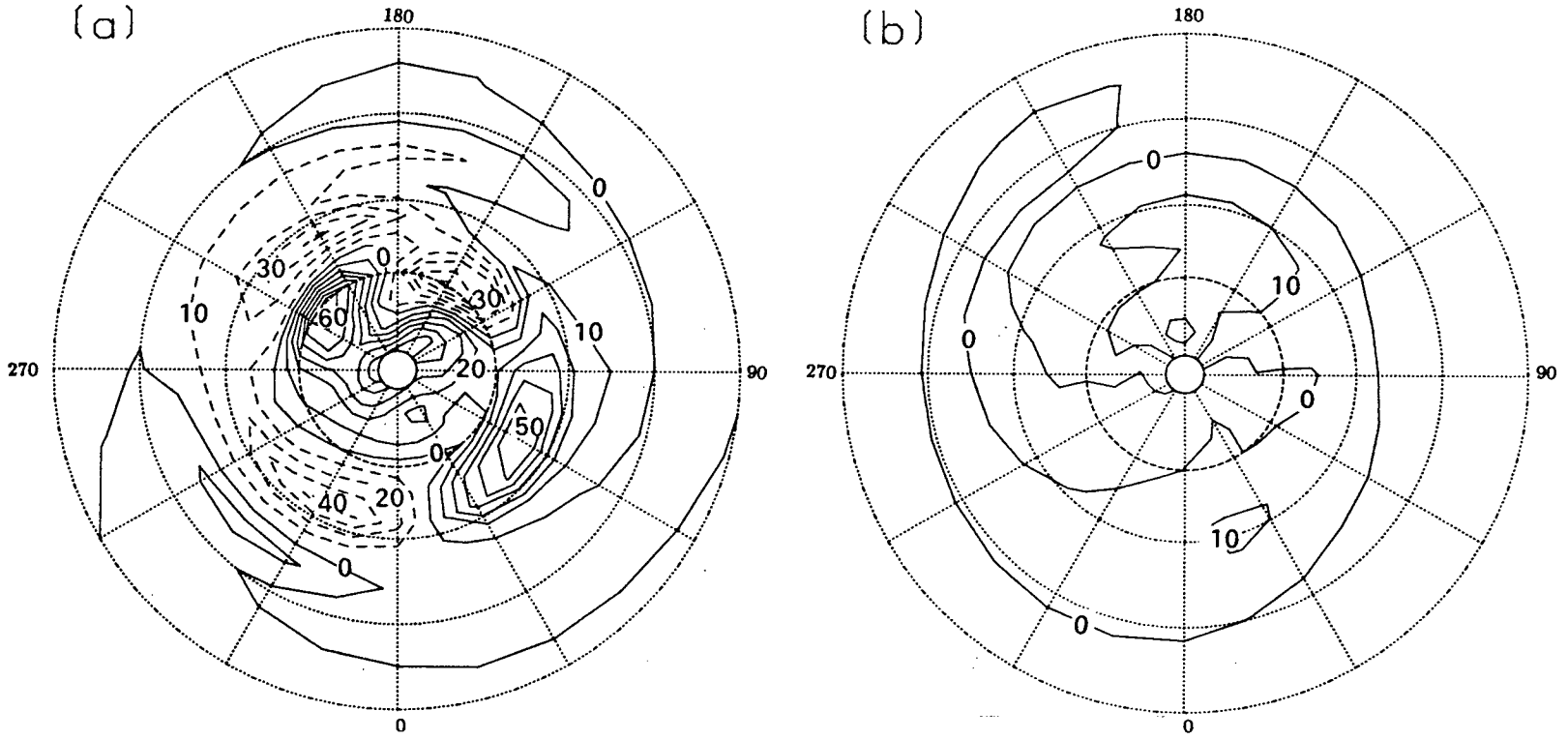


Figure 7.11. Forcing terms for the divergent mass-flow on January 25 1979 at $\eta = 6.7$ (a) $\partial\sigma_\eta/\partial t$ and (b) $\partial W/\partial\eta$, both in $kgm^{-2}day^{-1}$ with contour interval $10kgm^{-2}day^{-1}$ (note that this is ten times larger than in figure 7.1) between 4°N and 84°N.

stratopause the direct diabatic forcing of the potential vorticity is also more relevant (the radiative dissipation coefficient being about twice as strong as in the middle stratosphere - see chapter 8); however, the errors in the geostrophic potential vorticity are still larger than the contribution of the divergence to the budget.

This chapter must therefore end pessimistically, concluding that a complete study of the potential vorticity budget using satellite-derived dynamical quantities seems error-prone and perhaps impossible. However, some of the techniques discussed could perhaps be applied to the analysis of model data.

CHAPTER 8

RADIATIVE DISSIPATION OF TEMPERATURE PERTURBATIONS.

8.1. Introduction: Simple Expressions for Radiative Dissipation.

The analysis of chapter 2 showed that the zonal-mean atmospheric state can be maintained away from radiative equilibrium by the action of planetary waves; specifically, any radiative or mechanical dissipation or nonlinearity of these waves can lead to a northward flux of potential vorticity, thereby leading to an acceleration of the zonal-mean zonal flow (Dickinson, 1969; Andrews, 1985; Haynes and McIntyre, 1987). This maintenance of the atmosphere from radiative equilibrium leads to the existence of regions of net heating and cooling which in turn cause a diabatic circulation to exist (e.g. Fels, 1985), as has been discussed.

Whilst there is no doubt that nonlinearity plays a major role in the maintenance of the winter stratosphere from equilibrium, the waves must also be radiatively dissipated at such times because they cause large, zonal-asymmetries in the temperature field which, to a crude approximation, cool most strongly where they are warm and least strongly in cold regions; thus radiative processes tend to relax the structure back towards zonality. In the tropics the waves are thought to be linear and radiatively dissipated (e.g. Holton and Lindzen, 1972; Gray and Pyle, 1987).

It is thus of some importance to have estimates of the radiative dissipation timescales for planetary waves in the stratosphere; these are particularly useful for mechanistic models of wave propagation, in which the physical processes are treated very crudely so that attention can be devoted to the wave propagation in a simplified framework. Several estimates of these radiative dissipation timescales have been presented in the last two decades; none of these is entirely complete, since they consider only 'Newtonian' cooling, valid for infinitely deep disturbances to a standard atmospheric profile (Dickinson, 1973; Kiehl and Solomon, 1986), or idealised, small-amplitude, sinusoidal perturbations to a slowly-varying basic state (Fels, 1982): the scale-dependent dissipation rate. The photochemical acceleration of the radiative damping (Strobel, 1977; Hartmann, 1978) is often not included even though it has been shown to be important by Ghazi *et al.* (1979, 1985); Haigh (1985) has considered the scale-dependence of the photochemical acceleration

coefficient.

The purpose of this chapter is to present a set of observed dissipation coefficients for observed atmospheric perturbations; thus the effect of the vertical scale of the perturbations is implicitly included in the calculations; the photochemical contribution is also included. The relative contributions of CO_2, O_3 and H_2O to the dissipation rate are considered.

There seem to be two distinct approaches to calculating radiative dissipation coefficients. The first is to consider departures from an equilibrium temperature profile (either radiative or radiative-photochemical equilibrium); the second is to consider departures from observed temperature profiles. Comparison of the two approaches is illuminating. Denote the equilibrium temperature as T_e and the observed zonal-mean temperature by T_o; zonally asymmetric departures from these two fields are denoted as δT and T', so that the actual temperature may be expressed as:

$$T = T_o + T' = T_e + \delta T.$$

The radiative cooling rate, J_{LW}, may then be expanded about either mean temperature, so that:

$$J_{LW} = J_{LW}(T_o) + \left(\frac{\partial J_{LW}}{\partial T}\right)T' = J_o + \alpha T' \qquad (8.1)$$

$$= J_{LW}(T_e) + \left(\frac{\partial J_{LW}}{\partial T}\right)\delta T = J_e + \alpha_e \delta T$$

(where the higher order terms in the Taylor expansion have been neglected on the assumption that $T' << T_o$ and $\delta T << T_e$). The two coefficients, α and α_e, are the inverse relaxation times for disturbances to the observed and equilibrium states. These two coefficients can be related, since:

$$\alpha = \left(\frac{\partial J_{LW}}{\partial T}\right)_{T=To}$$

$$= \left(\frac{\partial J_{LW}}{\partial T}\right)_{T=Te} + \left(\frac{\partial^2 J}{\partial T^2}\right)_{T=Te}(T_o - T_e)$$

$$= \alpha_e + (T_o - T_e)\frac{\partial \alpha_e}{\partial T}$$

(assuming that $T_o - T_e << T_e$). Thus, assuming that α_e is only a slowly varying function of temperature and that the observed zonal-mean temperature does not depart too far from radiative equilibrium, these two dissipation coefficients should not be too different. Shine (1987) has shown that the latter assumption may not be valid in the polar night, where the large departures from the radiative equilibrium temperature field causes the dissipation time to be much greater than that predicted using the 'small perturbation' arguments given here.

Note that the physical difference between using α_e and α is that in the first case the entire atmosphere must be assumed to relax back to radiative equilibrium, whilst in the second case the assumption is that the zonally-asymmetric perturbations relax back to zero (and then there must be a different relaxation coefficient for the zonal-mean state).

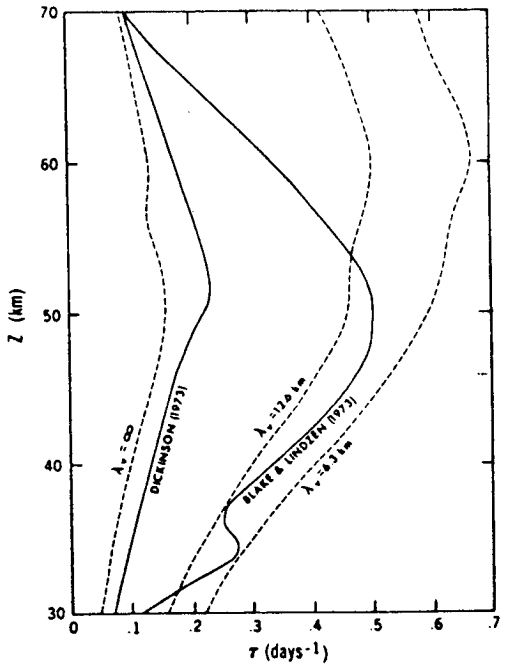


Figure 8.1. The inverse radiative dissipation times as a function of altitude calculated by Dickinson (1973) (solid line), Fels (1982) (dashed lines) for three vertical wavelengths (infinite, 12.6km and 6.3km) and Blake and Lindzen (1973). (From Fels, 1982).

8.2. Previous Estimates of Radiative Dissipation Coefficients.

The first useful estimates of so-called Newtonian cooling rates were presented for perturbations about an observed mid-latitude temperature profile by Dickinson (1973). Radiative cooling rates for infinitely deep, small-amplitude (0.1K) perturbations to this profile were calculated and the inverse dissipation

time as a function of height calculated; this so-called 'Newtonian' cooling coefficient varies from around $0.07day^{-1}$ at 30km to $0.22day^{-1}$ at 50km, above which it decreases slightly with increasing altitude (figure 8.1). Thus the relaxation time varies from more than 10 days at 30km to less than 5 days at 50km.

Kiehl and Solomon (1986) repeated the calculations of Dickinson, using observed zonal-mean LIMS data, to show that the Newtonian cooling time varies with latitude and season (this is predominantly because the temperature profile, T_o, varies with latitude and season). For January (figure 8.2) the relaxation time ranges from around 80 days ($\alpha=0.013day^{-1}$) at 20km in the tropical stratosphere to around 20 days ($\alpha=0.05day^{-1}$) in mid-latitudes; the dissipation coefficient decreases with height to less than 10 days ($\alpha>0.1day^{-1}$) at 55km. In middle latitudes these results agree quite well with those of Dickinson at low levels, but the inverse dissipation time of Kiehl and Solomon is rather smaller at high levels. These results emphasise the need to allow for these large latitudinal variations in relaxation time in simple models of planetary wave propagation.

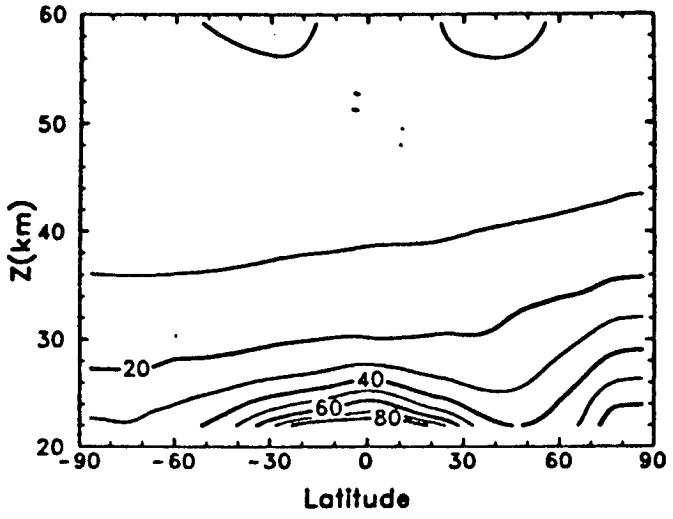


Figure 8.2. The latitude-height distribution of Newtonian cooling timescale (α^{-1}, days) calculated by Kiehl and Solomon (1986) for small (0.1K) perturbations from the zonal-mean LIMS data for January 1979.

A second class of problem was considered by Fels (1982). These arise because real atmospheric disturbances do not have a uniform vertical structure but tend to vary, in both amplitude and phase, with altitude. This is particularly relevant in the tropics, where the observed Kelvin waves have vertical wavelengths of less than 10km; in middle latitudes the waves are generally deeper than this. The problem is also relevant for gravity waves in the atmosphere, which have a vertical wavelength of less than 1km, but these are not resolved by satellite data (LIMS has a nominal vertical resolution of 1.8km).

The relevance of this vertical structure arises because radiative cooling is not dependent on the local temperature only, but also upon the state of the entire atmosphere (see chapter 4); a wavelike vertical structure can be dissipated radiatively by the exchange of heat between adjacent warm and cool levels. Furthermore, the rate of dissipation can be expected to vary with vertical wavelength because of variations in the optical depth of the atmosphere associated with the waves. Fels calculated the dissipation

timescale for idealised sinusoidal disturbances, with an amplitude of 1K; the dissipation times were presented as a function of vertical wavenumber. Cooling in the 15μm CO_2 and 9.6μm O_3 bands was considered separately. For CO_2 the inverse dissipation time at 60km was found to increase by a factor of four between an infinitely deep perturbation (the Newtonian cooling limit) and a wavelength of 4km (this latter wave would be dissipated in around 1.5days); at 20km the dissipation timescale also decreased (figure 8.3a). The contribution from the O_3 bands was weaker than that due to CO_2 (figure 8.3b), but showed a more complex relationship with wavelength; at 30km and 40km there was a fourfold increase in the dissipation rate between the infinitely deep and 4km vertical waves; however, at higher levels (50 and 60km) there was an initial decrease in dissipation time as the wavelength decreased, but then the atmosphere became transparent over a range of several wavelengths, so that the dissipation rate remained constant as the vertical wavelength decreased. The total dissipation rate as a function of height is shown in figure 8.1; differences between Fels' infinitely deep perturbation and that of Dickinson are most likely a consequence of the different radiation transfer models or temperature profiles used in the two studies. For shorter vertical wavelengths the dissipation times are considerably reduced; at 50km α is $0.5day^{-1}$ for a 12.6km vertical wavelength and $0.6day^{-1}$ for a 6.3km perturbation.

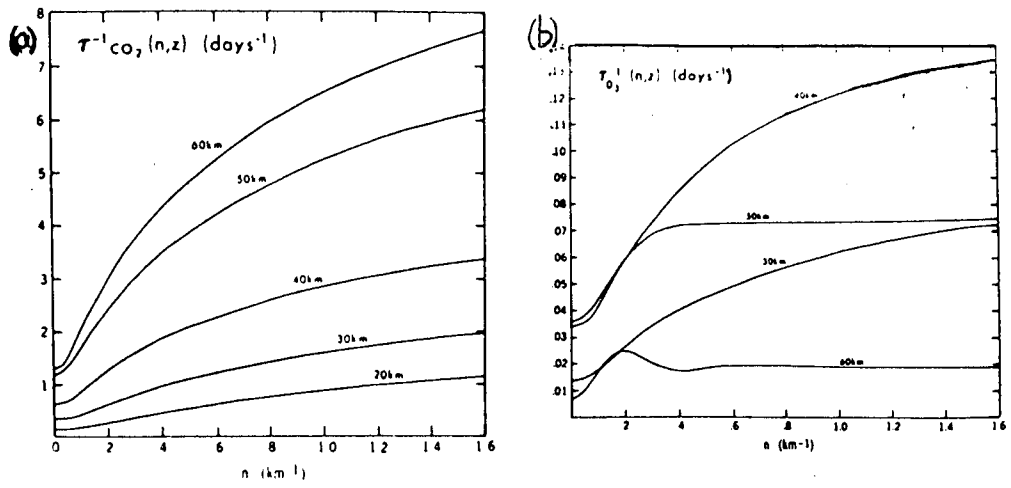


Figure 8.3. The inverse dissipation time, $\alpha = \tau^{-1}$, as a function of vertical wavenumber of a sinusoidal perturbation for (a) the 15μm bands of CO_2 and (b) the 9.6μm bands of O_3 (the ordinate on b is twice as high resolution as that on a), as calculated by Fels (1982).

The studies reviewed so far have included varying complexity in the form of the temperature perturbations, but have not allowed variations in the ozone distribution. This could be important because of the temperature dependence of the ozone mixing ratio; this is particularly true in the high stratosphere, where the photochemical relaxation time is short so that the ozone and temperature are often observed to be out of phase (e.g. Barnett *et al*, 1975; see also figure 5.8 of this thesis). Blake and Lindzen (1973) considered the importance of temperature dependent ozone variations in their calculation of dissipation coefficients for infinitely deep waves (figure 8.1); their photochemical model is now known to be invalid (e.g. Fels, 1982) but their calculations do serve to show that the dissipation time is sensitive to temperature-dependent ozone variations. Ghazi *et al.* (1979, 1985) used observations of the atmospheric temperature and ozone structure in the winter hemisphere to calculate dissipation coefficients in the mid-latitude stratosphere; hence their calculations implicitly include the temperature dependence of the ozone distribution. Temperature data from the daily NMC analyses were used with the SAGE ozone data; the longitudinal distribution of

radiative cooling rates were calculated using the radiation transfer model of Ramanathan (1976).

Ghazi *et al.* also calculated the photochemical acceleration of the radiative dissipation rate, which arises because the temperature and ozone distributions at high levels are out of phase; thus the strongest solar heating occurs at the longitude of minimum temperature, thereby dissipating the wave. Because of the temperature dependence of the ozone distribution the perturbation solar heating rate can be expressed in terms of the temperature perturbation, in the same way as the expression for the longwave cooling perturbation. Thus:

$$J'_{UV} = -\beta T', \tag{8.2}$$

the coefficient β being related to the temperature dependence of the ozone distribution (e.g. Hartmann, 1981). It should be realised that this is a rather simplified manner of writing the coefficient β, since the opacity effect means that the perturbation solar heating is dependent not only upon the local temperature perturbation, but also that at higher levels (Strobel, 1977; Hartmann, 1978).

Ghazi *et al.* used equations (8.1) and (8.2) to calculate α and β from the atmospheric data. By multiplying these expressions by T' and performing the zonal average, a diagnostic relationship for the radiative dissipation and photochemical acceleration times is obtained:

$$(\alpha,\beta) = -([J'_{LW}T'], [J'_{UV}T'])/[T'^2]. \tag{8.3}$$

The results obtained using this relationship showed a broad range of values at each height, so that the standard deviation about the mean value of α was quite large (figure 8.4). Below 45km the radiative dissipation coefficients agree well with the Newtonian cooling coefficients of Dickinson and Fels, but at higher levels the values of Ghazi *et al.* indicate dissipation times around twice as long as the others. The photochemical acceleration coefficients at high levels are of the same magnitude as the radiative dissipation coefficients, so that the inclusion of these in the radiative dissipation timescale gives results very similar to those of Dickinson, who considered only the longwave component. It is notable that the standard deviation in the photochemical

acceleration coefficients calculated by Ghazi *et al.* tends to be the same size as the coefficients themselves.

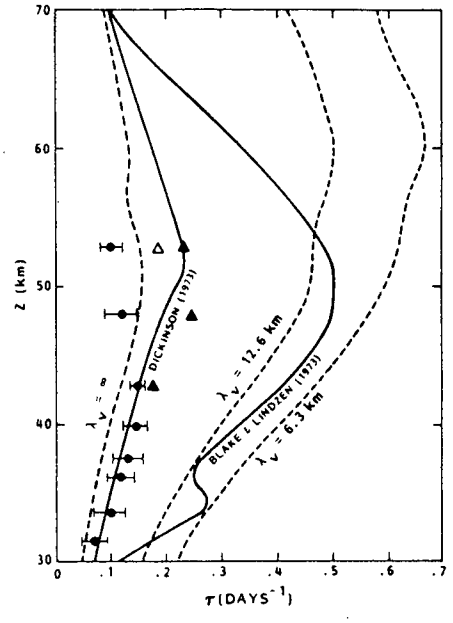


Figure 8.4. As figure 8.1 but with the dissipation coefficients of Ghazi *et al.* near 55° in the winter hemisphere added; solid circles are the mean long wave dissipation coefficients (with standard deviation shown); solid triangles are the total dissipation coefficients, including the effect of photochemical acceleration.

Haigh (1985) considered the scale dependence of the photochemical acceleration rate due to solar absorption; the response of the calculations to the inclusion of 9.6μm band cooling in the radiative-photochemical equilibrium state and the temperature-dependent variations in ozone mixing ratio were considered in this study. Both effects were found to be important, as is evident in figure 8.5; for an infinitely deep 1K perturbation the photochemical acceleration coefficient was decreased by around 0.01days^{-1} between 3-0.5mbar when the 9.6μm radiation transfer was included; a similar decrease was obtained for the case of a 7km vertical wavelength, when the photochemical acceleration of the relaxation coefficient tended to be around twice as strong as for the infinitely deep perturbation.

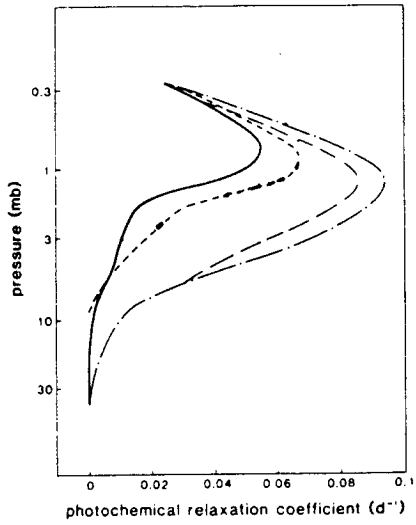


Figure 8.5. Photochemical relaxation coefficients calculated by Haigh (1985) for: - - -, a uniform perturbation with no 9.6μm coupling; ___, a uniform perturbation with 9.6μm coupling; -·-·-, short wavelength perturbation, no 9.6μm coupling; __ __, short wavelength perturbation, with 9.6μm coupling.

In the next section the calculations of Ghazi *et al.* (1985) are repeated using the LIMS data. The contributions of CO_2, O_3 and H_2O will be considered separately so that the importance of these different gases and their relationship to the temperature distribution can be considered. Data for tropical profiles are considered in section 8.4.

8.3. Calculation of Dissipation Rates: Mid-Latitude Winter.

The method of Ghazi *et al*, described in the previous section, has been used to calculate α and β from the LIMS temperature, ozone and water vapour data and their derived heating rates, which were calculated at 24 longitudes. As well as the total longwave dissipation coefficient, α, its three components due to CO_2, O_3 and H_2O have been calculated: their relative importance at different heights will be discussed. Rather than presenting Fourier analysed results the net dissipation coefficient for the observed perturbations is presented; this saves making a somewhat artificial partitioning of the data when the atmospheric state is highly nonlinear; this could be important because the scale-dependence of the results is a non-linear function of the vertical temperature structure. In practice, the Fourier analysed coefficients do not show significant variations between wavenumbers (at least for the three longest planetary waves), particularly when the standard deviation of α for each of the waves is considered (e.g. Ghazi *et al*, 1985). Data from five days

at 60° in the winter hemisphere are considered; these are 5 November 1978, 20 December 1978, 25 January 1979 and 22 February 1979 at 60°N and 25 May 1979 at 60°S. The first and last of these represent early winter profiles for each hemisphere, the second and third are for strong wavenumber-1 disturbances in the middle stratosphere and the fourth for a strong wavenumber-2 perturbation in the middle stratosphere; table 8.1 shows $[T'^2]$ as a function of height for each of these days, providing an indication of the strength of the perturbations. The strongest disturbances are evident at 60°N on 25 January, whilst the early winter days generally show the weakest non-zonalities.

Table 8.1. The zonal-average of the square of the temperature perturbations ($[T'^2]$) at 60° in the winter hemisphere as a function of height for the five days analysed.

Approx Ht (km).	5 Nov.	25 May.	20 Dec.	25 Jan.	22 Feb.
55.0	34	9	22	190	58
51.5	28	17	62	98	89
48.0	24	20	101	21	158
44.5	38	20	96	70	180
41.0	66	30	84	125	125
37.5	66	20	38	194	74
34.0	72	27	22	270	55
30.5	40	26	19	304	50
27.0	16	23	10	314	87
23.5	8	18	11	286	106
20.0	5	13	11	212	97

The radiative dissipation coefficients are tabulated as a function of approximate altitude (table 8.2) for these five days; the mean and standard deviation are also shown. Below 48km there is little scatter in the results; however, at higher levels there is more variation in the daily calculations. This variability is consistent with the results of Ghazi *et al.* (1985), but the mean values are not; the results presented here agree much more closely at high levels with those of Dickinson (1973) and Fels (1982) than those of Ghazi *et al* (figures 8.6 and 8.4). This will be discussed later.

Table 8.2. Radiative Dissipation Coefficient, α, for five days in the mid-latitude winter (60°), their mean and standard deviation.

Approx Ht (km).	5 Nov.	25 May.	20 Dec.	25 Jan.	22 Feb.	Mean.	Std Dev.
55.0	0.26	0.25	0.16	0.24	0.09	0.20	0.03
51.5	0.24	0.18	0.18	0.28	0.13	0.20	0.02
48.0	0.20	0.17	0.18	0.20	0.16	0.18	0.01
44.5	0.13	0.13	0.17	0.13	0.17	0.15	0.01
41.0	0.11	0.10	0.15	0.11	0.15	0.12	0.01
37.5	0.09	0.06	0.11	0.09	0.12	0.09	0.01
34.0	0.09	0.06	0.10	0.08	0.11	0.09	0.01
30.5	0.07	0.05	0.09	0.06	0.07	0.07	0.01
27.0	0.04	0.04	0.05	0.05	0.06	0.05	0.01
23.5	0.03	0.03	0.04	0.04	0.05	0.04	0.00
20.0	0.02	0.02	0.03	0.02	0.03	0.02	0.00

The close agreement with the Newtonian cooling calculations of Dickinson (1973) and Fels (1982) reinforces the notion that scale dependence is unimportant for the radiative dissipation of observed mid-latitude perturbations - at least when the data are zonally averaged. This occurs because the vertical wavelength is larger than the depth over which the atmospheric absorption is unity in the relevant bands. The results for an infinitely deep, 1K perturbation to the observed temperature profile at 60°N on 25 January 1979 are also shown on figure 8.6; these reaffirm this notion, since they are very similar to those calculated for the observed perturbations.

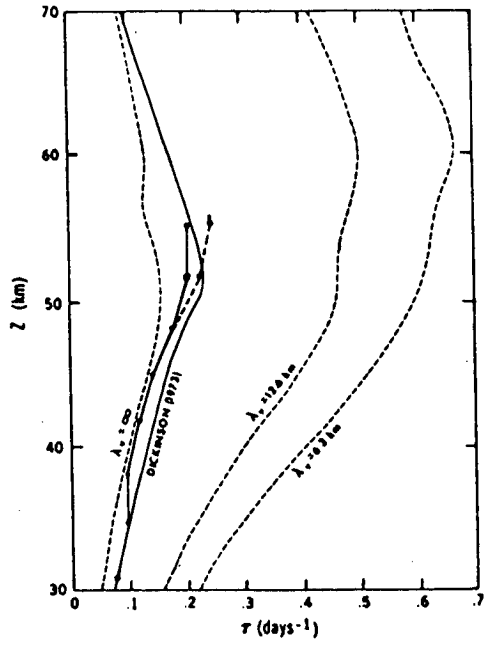


Figure 8.6. Radiative dissipation coefficients as a function of height, as in figure 8.1, but with the mean thermal dissipation coefficients, α (solid line), and total dissipation, $\alpha+\beta$ (dashed line), coefficients for 60° in the winter hemisphere calculated in this study superimposed.

It is of some interest to determine the contribution of each of the three gases to the longwave dissipation coefficient. Data from 25 January 1979 are discussed. The temperature, ozone and water vapour distributions and the associated heating and cooling rates as a function of longitude and height at 60°N are shown in figure 8.7. The strong zonal asymmetries in the temperature and tracer distributions, which are evident in the entire stratosphere, are also evident in the associated heating and cooling distributions. Maxima in the CO_2 cooling and temperature distributions are coincident; similarly, peaks and troughs in the ozone and water vapour distributions are echoed in their cooling distributions whilst maximum solar heating rates at any height coincide with ozone maxima.

The relative contributions to the dissipation timescale vary with height (table 8.3). At the highest levels α_{CO2} accounts for around 80% of the total dissipation, most of the remainder being due to ozone, with only a minor contribution from water vapour. In the layer 37.5–44.5km water vapour contributes around half as much to the cooling as O_3, but the CO_2 contribution is still dominant. The situation is somewhat different in the lower stratosphere, where α_{CO2} causes only half of the total dissipation. This result is anticipated from the calculations of chapter 4, which indicate that ozone and water vapour

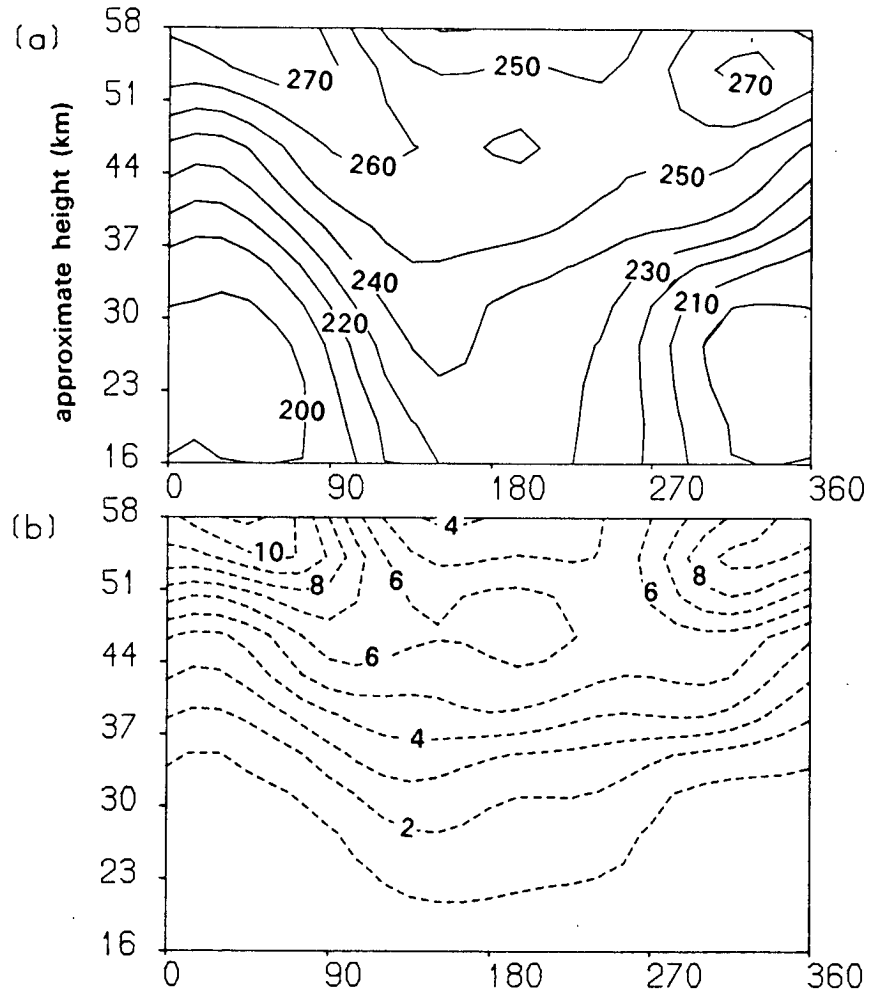


Figure 8.7. Longitude-height distributions, at 60°N on 25 January 1979, of (a) temperature (K, contour interval 10K); (b) Cooling due to the 15μm bands of CO_2 ($Kday^{-1}$, contour interval $1Kday^{-1}$). (Figure continues.)

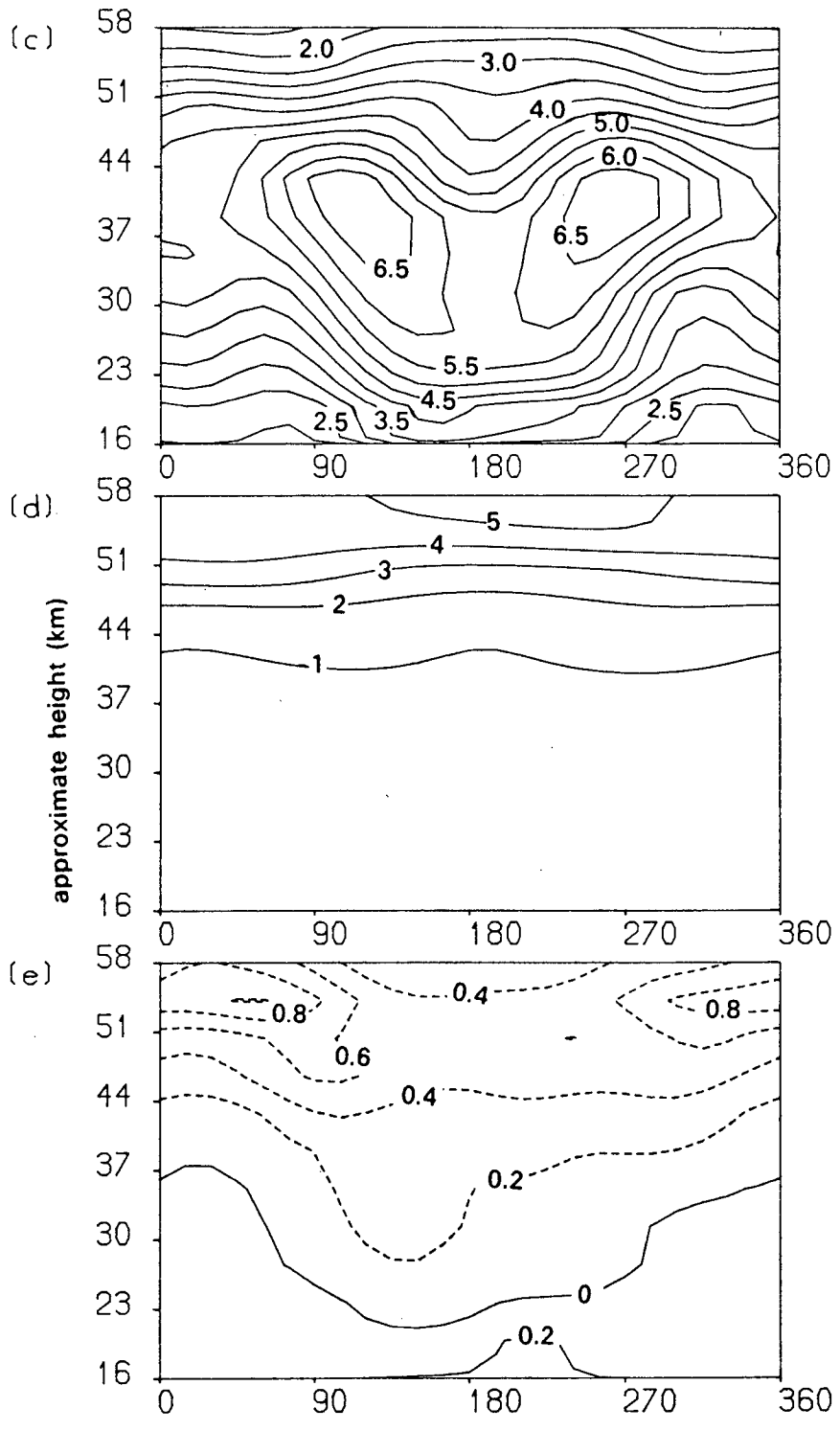


Figure 8.7 (continued). (c) ozone (ppmv, contour interval 0.5ppmv); (d) solar heating rate ($Kday^{-1}$, contour interval $1Kday^{-1}$); (e) cooling due to the 9.6μm bands of ozone ($Kday^{-1}$, contour interval $0.2Kday^{-1}$ in regions of heating – solid contours, and $0.5Kday^{-1}$ in regions of cooling – dashed contours). (Figure continues.)

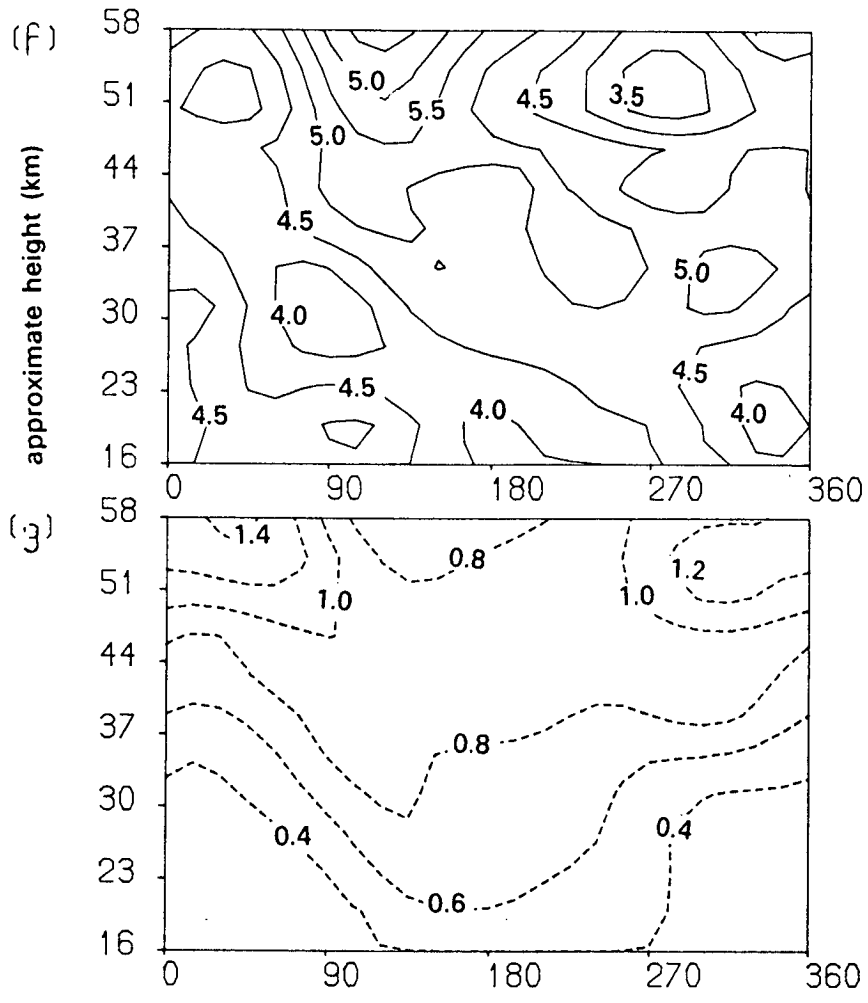


Figure 8.7 (continued). (f) water vapour (ppmv, contour interval 0.5ppmv); (g) cooling due to water vapour (Kday^{-1}, contour interval 0.5Kday^{-1}).

play a more important role in the radiative balance of the low stratosphere than at high levels. At 20km ozone caused a slight acceleration (i.e. α_{O3} was negative); this was insignificant at this latitude.

Table 8.3. Inverse dissipation times, α (day^{-1}), as a function of altitude due to the different gases on 25 January 1979 at 60°N.

Approx Ht (km).	CO_2	O_3	H_2O
55.0	0.18	0.04	0.02
51.5	0.20	0.05	0.02
48.0	0.14	0.04	0.01
44.5	0.10	0.02	0.01
41.0	0.08	0.02	0.01
37.5	0.07	0.02	0.01
34.0	0.05	0.02	0.01
30.5	0.04	0.02	0.01
27.0	0.03	0.01	0.01
23.5	0.02	0.01	0.01
20.0	0.01	0.00	0.01

Above 45km the photochemical acceleration coefficient was found to contribute significantly to the radiative dissipation of planetary waves by Ghazi *et al.* (1985); they found that β was around $0.1day^{-1}$ near 50km, which was comparable with their radiative dissipation coefficient. Our calculations do not show such a strong photochemical acceleration coefficient (table 8.4); at 50 and 55km values of 0.03 and $0.04day^{-1}$ were calculated, which are only half as strong as the results of Ghazi *et al.* at a similar latitude. There is, however, a considerable dispersion about this mean value: on 22 February 1979 the photochemical acceleration coefficient was twice the average. Between 45 and 50km the photochemical acceleration coefficient was negligible in comparison to the infrared dissipation timescale; at these lower levels it was often negative, indicating acceleration of the mean flow, a consequence of the maxima in solar heating being coincident with the temperature maxima at these levels.

Table 8.4. Photochemical acceleration of the radiative dissipation coefficient, β (day^{-1}), for five days at 60°N in winter.

Approx Ht (km).	5 Nov.	25 May.	20 Dec.	25 Jan.	22 Feb.	Mean.	Std. Dev.
55.0	0.04	0.02	0.03	0.04	0.08	0.04	0.01
51.5	0.03	0.01	0.02	-0.00	0.08	0.03	0.01
48.0	0.00	-0.01	-0.00	-0.01	0.01	-0.00	0.00
44.5	0.00	-0.00	-0.00	0.01	-0.00	0.00	0.00

Thus, our calculations differ from those of Ghazi *et al.* in two respects: our longwave dissipation coefficient is stronger and our solar contribution is weaker. The net result is that our total dissipation time at high levels is slightly shorter than that calculated by Ghazi *et al.*: this is probably a coincidence. The reason for these differences is unclear; they could arise because of the use of different data sets or radiation transfer models; since both carbon dioxide radiation transfer models were based on that of Ramanathan (1976) the different temperature analyses could cause for the differences in α; the discrepancy in β is more likely to be caused by the radiation models used: that of Lacis and Hansen (1974) employed by Ghazi *et al.* does not extend above 55km, so that variations in the ozone distribution above this level are not accounted for, whereas the model used in these calculations (Strobel, 1978) takes full account of the ozone variations to higher levels.

8.4. Calculation of Dissipation Rates: Tropical Profiles.

Whereas in middle latitudes the observed perturbations are generally deep in structure, those detected in tropical data are not. Tropical waves in LIMS data, reported by Salby *et al.* (1984) had vertical wavelengths of 10-40km; the results of Fels (1982) suggest that the scale of such perturbations may be sufficiently small for the radiative dissipation rate to be significantly different from the 'Newtonian' cooling rate calculated by Kiehl and Solomon (1986), shown in figure 8.2. Indeed, the dissipation rates calculated by Kiehl and Solomon showed a dissipation timescale ranging from 200 days near the

tropopause to 10 days at higher levels in the tropical stratosphere, whereas the Holton and Lindzen (1972) model of the quasi-biennial oscillation requires a dissipation time of around three days near 40km to give an oscillation of the correct period and structure.

The waves detected by Salby *et al.* were eastward propagating and were of wavenumber-1 or 2; a slow (phase speed near $10ms^{-1}$) and a fast (phase speeds of $\sim 120ms^{-1}$, wavenumber-1 and $\sim 60ms^{-1}$, wavenumber-2) component of each wavenumber were detected, the fast component being dominant in the upper stratosphere. Holton and Lindzen showed that Kelvin waves can provide the momentum source for the westerly phase of the quasi-biennial oscillation; successive enhancements of their model have verified their conclusions, and indicated that a similar mechanism can also account for the westerly phase of the semi-annual oscillation of the zonal-mean wind at higher levels of the middle atmosphere. Coy and Hitchman (1984) showed that the Kelvin waves observed by Salby *et al.* could provide the momentum source for the semi-annual oscillation; they used a simple model, in which the waves were thermally dissipated by a coefficient which increased linearly from zero at 27km to $1day^{-1}$ at 70km. This simple parametrisation prompts speculation as to the actual dissipation coefficients in the tropical stratosphere; the purpose of this section is to calculate α and β for the observed tropical temperature profiles.

Profiles from 4°N or 4°S for the same five days analysed in the previous section are discussed. In general, the thermal dissipation coefficients are stronger than in the mid-latitude case. Near 50km the mean α is $0.29day^{-1}$ (table 8.5), with individual values ranging from $0.18day^{-1}$ to $0.42day^{-1}$. Near 30km α is of the order of $0.15day^{-1}$, whereas in the low stratosphere it is much smaller. However, the mean value of around 0.0 at 20km is misleading, as it is the average of four weakly dissipating cases and one strongly forcing case (α=-0.28 at 4°N on 5 November 1978). Indeed, in the region 20-25km there is often evidence of α being negative; this arises because α_{O3} is negative in this region of heating in the 9.6μm bands, and on occasions it dominates the dissipation by CO_2 and H_2O. This suggests that Kelvin waves can often be amplified in the low stratosphere, on a timescale of 4-10 days.

Table 8.5. The radiative dissipation coefficient, α (day^{-1}), for five days, their mean and standard deviation, at 4° in the winter hemisphere.

Approx. Ht (km)	5 Nov.	20 Dec.	25 Jan.	22 Feb.	25 May.	Mean.	Std. Dev.
55.0	0.26	0.25	0.26	0.25	0.27	0.26	0.00
51.5	0.18	0.21	0.29	0.42	0.35	0.29	0.04
48.0	0.29	0.30	0.33	0.22	0.30	0.28	0.02
44.5	0.24	0.28	0.29	0.28	0.26	0.27	0.02
41.0	0.22	0.27	0.26	0.25	0.18	0.24	0.02
37.5	0.21	0.23	0.22	0.20	0.17	0.21	0.01
34.0	0.21	0.16	0.19	0.17	0.17	0.18	0.01
30.5	0.17	0.17	0.13	0.13	0.13	0.15	0.01
27.0	0.05	0.13	0.05	0.28	-0.08	0.09	0.06
23.5	-0.05	-0.05	-0.02	0.23	-0.05	0.01	0.05
20.0	-0.28	0.02	0.07	0.10	0.07	0.00	0.06

Thus there seems to be persistent evidence that waves are forced in the low tropical stratosphere. In mid-latitudes of the winter hemisphere, discussed in the previous section, it was found that because ozone absorbs radiation emitted in the troposphere it tends to reduce the radiative dissipation coefficient, since this absorption is positively correlated with the temperature distribution, making α_{O3} negative. This is also the case in the tropical lower stratosphere. However, the fundamental difference between the mid-latitude and tropical case occurs because there is much less radiative emission by the bands of CO_2 and H_2O in the cold equatorial regions. The consequence of this is that α_{CO2} and α_{H2O}, although positive, are weak in the tropics, and the ozone-induced amplification is dominant, causing an overall enhancement zonal-asymmetries in the temperature structure, or amplifying the waves, in the lower stratosphere.

The overall importance of this acceleration is not yet clear; it is generally fairly weak and can only exist when a temperature perturbation already exists, so it does not appear to be a mechanism for wave generation, merely enhancement. Nevertheless, it warrants further attention (spectral analysis reveals that it can occur for both waves 1 and 2). A further point is the importance of clouds: inclusion of observed tropical clouds should be an

important feature of future studies, given the sensitivity of the stratospheric heating rate below 30km to their presence (Haigh, 1984; Gille and Lyjak, 1986; chapter 4 of this thesis). These calculations should be repeated with observed tropospheric cloud distributions to check their validity.

In an attempt to provide some justification of the robustness of the result, monthly-mean values of the dissipation coefficients have been calculated from the monthly-averaged, three-dimensional fields of temperature, ozone and water vapour. These fields should provide more accurate results in the sense that any short-term fluctuations, including noise, in the data are smoothed by the averaging process; however, they are still sensitive to possible errors in the upward radiant energy flux from the troposphere. These data reveal (table 8.6) that the dissipation coefficients are negative in the low stratosphere (at 4° in the winter hemisphere); the feature was also evident at 8° in each month, but not at the equator itself. It was also insensitive to variations in the zonal-mean cloud amount: repeating the calculations with the zonal-mean climatological cloud distribution or no cloud at all gave no significant change (<1%) in the calculated value of α.

Table 8.6. Monthly-averaged thermal dissipation coefficients, α (day^{-1}), at 4°N.

Approx Ht (km)	Dec.	Jan.	Feb.
55.0	0.25	0.24	0.33
51.5	0.24	0.21	0.27
48.0	0.23	0.29	0.27
44.5	0.20	0.24	0.25
41.0	0.21	0.22	0.24
37.5	0.19	0.22	0.20
34.0	0.17	0.14	0.19
30.5	0.15	0.15	0.07
27.0	-0.12	0.16	-0.05
23.5	-0.05	-0.06	0.06
20.0	0.01	0.06	-0.05

The dissipation coefficients are now compared with the previous estimates. They fall between the 'Newtonian' cooling and scale-dependent calculations of

Fels (1982) (figure 8.8). Between 30–40km our calculation show that α is around $0.05day^{-1}$ less than Fels' result for a 12.6km vertical wavelength, whilst above 40km the difference exceeds $0.1day^{-1}$. There are several possible explanations for this; firstly, the observed stratospheric perturbations are generally quite small (often less than the 1K used by Fels); secondly, Fels used a mid-latitude temperature profile in his calculations whereas our results are valid for perturbations to a tropical temperature profile; thirdly, Fels did not allow the ozone to vary with the temperature, so that at high levels his dissipation coefficients are likely to be overestimated because generally ozone minima tend to coincide with temperature maxima near the stratopause; fourthly, Salby *et al.* noted that whereas the dominant waves in the low tropical stratosphere had vertical wavelengths of around 10km, those at higher levels were around 40km, so that the effect of scale dependence is likely to be reduced here. Further, the large daily variations (tables 8.5 and 8.6), which are associated with the variations in the daily wave structure (Salby *et al.*, Hitchman and Leovy, 1986) indicate that on occasions the dissipation rate does reach that calculated by Fels, whilst at other times it is much smaller.

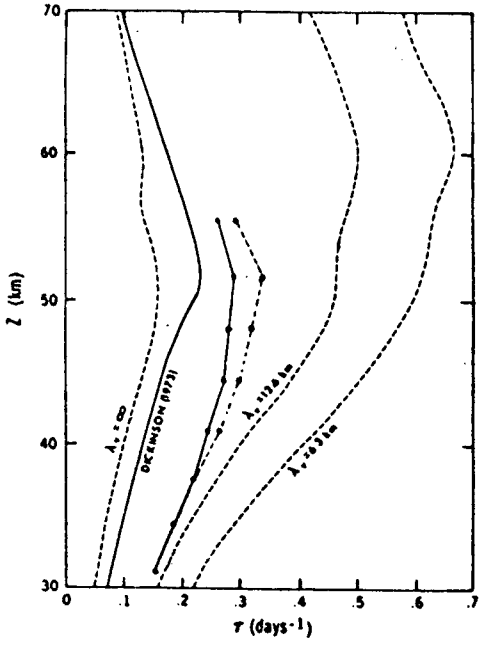


Figure 8.8. As in figure 8.1, but with the radiative dissipation coefficients, α (solid line) and $\alpha+\beta$ (dashed line), averaged over the five daily tropical profiles discussed in the text included.

The photochemical acceleration coefficient, β, showed much variation for each of the five days of the analysis; it was often negative and generally quite

small, although at high levels it exceeded $0.05day^{-1}$ on several occasions; the means tend to be smaller than the daily values would suggest because of the frequent occurrence of forcing (table 8.7). The reason for these negative values of β seems unclear; they also appear quite frequently in the monthly-mean calculations, so it seems unlikely that they are due to noise in the daily observations. Generally, and in the mean, β is positive (table 8.7) so that the dissipation rate is increased by the photochemical effect.

Table 8.7. The photochemical acceleration coefficient, β, for five days, and their mean, at 4° in the winter hemisphere.

Approx. Ht	5 Nov.	20 Dec.	25 Jan.	22 Feb.	25 May.	Mean.
55.0	-0.02	0.05	0.05	-0.03	0.11	0.03
51.5	0.02	0.07	0.07	-0.00	0.10	0.05
48.0	-0.01	-0.05	0.02	-0.01	0.12	0.01
44.5	0.11	-0.01	-0.02	0.04	0.04	0.03
41.0	0.05	0.03	0.03	0.00	-0.00	0.02
37.5	0.01	0.01	0.01	-0.01	-0.00	0.00
34.0	0.03	0.04	-0.00	0.00	-0.01	0.01
30.5	0.03	0.02	-0.03	-0.02	-0.02	0.00
27.0	0.01	-0.02	-0.03	-0.01	-0.01	-0.01
23.5	-0.02	-0.03	-0.04	-0.01	-0.03	-0.03
20.0	0.02	0.01	-0.01	-0.00	-0.00	-0.00

CHAPTER 9

CONCLUSIONS AND DISCUSSION.

In this section the major conclusions of this thesis are summarised and possible extensions to the work discussed.

The most fundamental aspect of this thesis has been the calculation of radiative heating and cooling rates and their combination to provide three-dimensional distributions of net radiative heating rate, which are conveniently presented in isentropic coordinates. These net radiative heating rates are the first component of the diabatic circulation; the horizontal part has been calculated using the steady state continuity equation. The diabatic circulation has been calculated in three dimensions, as well as in the zonal-mean case, where it compares favourably with the wide range of results obtained in previous studies. The results of this thesis are restricted to one seven month period, so that no measure of the interannual variability or the climatological situation is contained in them; future studies should be directed towards these calculations, once the necessary data becomes available.

In chapter 4 the radiative heating and cooling rates were discussed. Zonal-mean cross-sections of heating, calculated using the method of Strobel (1978), agreed reasonably well with other recent studies, although they were generally too strong near the stratopause and slightly weak in the mid-stratosphere. The previous studies show a broad range of results; our calculations did not appreciably extend the range of these results. Similarly, zonal-mean cooling rates calculated with the model of Haigh (1984) agree reasonably well with previous calculations. The linearity of the zonal-mean calculations was discussed: the fields calculated from the three-dimensional temperature and ozone data did not differ appreciably from the zonal-mean calculations (the worst errors of around 5% occurred near the stratopause). The sensitivity of the results to errors in the data was estimated; the largest uncertainties were in the low tropical stratosphere, arising because of the use of a zonal-mean cloud climatology which affects the radiation transfer in the 9.6μm band of O_3, and near the stratopause, where possible errors in the ozone distribution affect the calculated solar heating rates. Error bars on the solar heating rates reach 20% near the stratopause but are considerably less at lower levels at all latitudes; the terrestrial component has an accuracy of

around 15% at most levels. These cannot simply be added to obtain the error in the net heating rate because both results depend upon the ozone distribution; there is some cancellation, making the net heating rate accurate to around 20%.

Chapter 5 contains descriptions of the net radiative heating fields which form the basis of the calculations in chapters 6 and 7. Zonal-mean cross-sections agree qualitatively well with other recent studies; quantitative analysis reveals that they fall within the broad range of results calculated by previous authors, not all of whom impose an entirely consistent global constraint on the net radiation balance. Of the reviewed studies, only Gille *et al.* (1987) allow for latitudinal variations of static stability and for the mean northward heat transport in their adjustment of the heating fields; these effects were most important at high altitudes. The constraint imposed in this study, that there should be no globally-averaged mass transport through any isentropic level, was shown to be accurate to within 1%, which is better than the errors in the data themselves. Note that an additional assumption in this thesis is that only net radiative heating rates can transport mass across isentropic surfaces; this is a good assumption in the stratosphere, but in the mesosphere (and even near the tropopause) gravity wave-breaking causes nonlinear mixing of the fluid, so that it can transport mass across isentropic levels; in such cases the constraint should be generalised to include this non-diabatic component of cross-isentropic transport, since it will affect the net radiative energy budget of the middle atmosphere as well as the transport of trace gases.

Three-dimensional radiative heating fields were discussed in relation to the vortex structure of the northern hemisphere stratosphere at two levels, one near 30km in the mid-stratosphere, the other near the stratopause (50km). At the higher level the net cooling in the polar night was stronger than at the low level; in terms of vertical transport this means that matter can be transported downwards more rapidly at high levels. This cooling tended to be strongest on the eastern flank of the displaced polar vortex (regarded as the position of the maximum in potential vorticity), which is consistent with the advection of warm air from the tropics by the cyclonic flow.

From the three-dimensional distributions of radiative heating rate the mass-weighted zonal-mean vertical velocity was determined; this was then

used in the continuity equation to calculate the northward flow at each isentropic level: together, these give the zonal-mean diabatic circulation. There is an additional, transient, component of the mean-meridional-circulation, associated with the isentropic redistribution of mass by transient waves; although important on a daily basis this is negligible in the monthly-mean, so that the diabatic circulation is a very good approximation to the mean meridional circulation on such timescales. Zonal-mean models of the middle atmosphere can be expected to reproduce the slowly-varying, monthly-mean tracer distributions (Plumb and Mahlman, 1987), so that the zonally-averaged diabatic circulation is the relevant component of the mean meridional-circulation in such models. A comparison of the cross-isentropic transport of trace gases by the mean and eddy components of the diabatic flow was made; this showed that the mean advection was generally at least an order of magnitude stronger than the eddy component. The only exception occurred near the ozone maximum, where the mean vertical gradient vanishes. Thus, to a very good approximation, the cross-isentropic eddy transport may be neglected in zonal-mean modelling studies. The implication is that if K-theory is used to parametrise the eddy transport in terms of the zonal-mean tracer distribution only one K is needed: the meridional diffusion term. This may be determined from the dynamics of the middle atmosphere, since it is proportional to the northward eddy flux of potential vorticity (Tung, 1986). This provides a much simplified dynamical framework for zonal-mean tracer transport models, although it does not assist the calculation of radiative and photochemical distributions. Future research should be directed towards developing an isentropic-coordinate tracer transport model for middle atmospheric studies; observational and theoretical estimates of the diffusion coefficient are also required (and ultimately these should be calculated internally to the model rather than being specified externally). Although such models cannot predict the daily evolution of the middle atmosphere, they are useful for examining the response of the ozone distribution to anthropogenic perturbations in the carbon and chlorine distributions in the atmosphere; as it becomes feasible to perform full three-dimensional integrations of dynamical-radiative-photochemical models the zonal-mean model can still play a role in determining which of the many possible experiments it is worth performing in a three-dimensional model.

The three-dimensional diabatic circulation was calculated in chapter 7. At

high levels it was found to constitute a significant (up to 20%) portion of the horizontal flow in regions of small Rossby number, so that the flow is essentially non-geostrophic. At lower levels the relationship to the geostrophic flow was not so clear, since maxima in the diabatic northward flow did not coincide with minima in the geostrophic flow. On shorter timescales (such as daily fields) the transient component of the flow is important; this cannot be so reliably determined from a noisy data set. Furthermore, attempting to use these data to examine the potential vorticity budget is not possible because the noise in the calculated vorticity fields is of a similar magnitude to the various terms in the potential vorticity budget equation (advection by the non-geostrophic, including vertical, flow). Modelling studies could be employed to examine the vorticity budget, in the hope that advances in technology in the next two decades can lead to data sets of sufficient accuracy for practical application: requirements are a higher signal-noise ratio and better longitudinal resolution whilst retaining the vertical resolution of the LIMS data. Models could also be used to test the accuracy of the technique used in chapter 7 to calculate the divergent flow.

Because one of the mechanisms responsible for forcing the atmosphere from radiative equilibrium is the radiative dissipation of planetary waves it is essential that this process should be understood. Radiative dissipation coefficients presented in chapter 8 show that in mid-latitudes of the winter hemisphere the planetary wave damping is well approximated by the Newtonian cooling coefficient, but in the tropics and summer hemisphere this is not the case. The small vertical scale of the temperature perturbations in the tropics enhances the dissipation rate, since radiation transfer between levels of the atmosphere is important in such cases. In the low tropical stratosphere the zonal-asymmetry of the heating in the 9.6μm bands of ozone dominates the cooling by carbon dioxide and water vapour, resulting in an enhancement of the temperature waves; further investigation is required into the validity of this result, since it occurs in a region where the ozone data are possibly erroneous and where there is most likelihood of error because of the tropospheric structure (particularly cloud distributions) used in the study.

To summarise, this thesis has considered some aspects of radiative-dynamical interactions in the middle atmosphere; radiative dissipation times for planetary waves have been calculated, emphasising the importance of

photochemical and scale dependence to the total dissipation rate. The diabatic circulations in two and three dimensions have been examined The unimportance of the diabatic eddy terms in zonal-mean tracer transport models has been discussed, which enables considerable simplifications to be made in their formulation. The importance of the non-geostrophic, three-dimensional diabatic circulation, particularly at high levels in the winter stratosphere has been noted.

I. Acronyms and Symbols.

I.I. Acronyms.

CIRA	COSPAR International Reference Atmosphere.
COSPAR	Committee On Space, Planetary and Atmospheric Research.
FGGE	First GARP Global Experiment.
GARP	Global Atmospheric Research Program.
LIMS	Limb Infrared Monitor of the Stratosphere.
MAP	Middle Atmosphere Program.
MAT	Mapped Archival Tape (LIMS data).
NMC	National Meteorological Centre (U.S.A.).
SAGE	Stratospheric Aerosol an Gas Experiment.
SAMS	Stratospheric and Mesospheric Sounder.
SBUV	Solar Backscattered Ultraviolet Radiometer.
ppmv	parts per million by volume (10^{-6}).
ppbv	parts per billion by volume (10^{-9}).

I.II. Glossary of Symbols.

a	radius of Earth.
c	phase speed of wave; speed of light.
c_p	specific heat capacity at constant pressure.
f	planetary vorticity.
g	gravitational acceleration ($9.81 ms^{-2}$).
h	Planck's constant.
i	longitudinal index for finite differencing.
j	latitudinal index.

k	vertical index; zonal wavenumber of wave.
k_ν	monochromatic absorption coefficient.
p	pressure.
p_o	standard reference pressure (1000mbar).
q	quasigeostrophic potential vorticity.
r	dummy variable.
s	dummy variable.
t	time.
$\mathbf{u}=(u,v)$	horizontal velocity.
$\mathbf{u}_G$	geostrophic horizontal velocity.
$\mathbf{u}_\chi=\mathbf{U}_D/\sigma_\eta$	quasi-divergent horizontal velocity.
$w=D\eta/Dt$	vertical 'velocity' (rate of change of entropy).
$y=a\phi$	northward distance.
z	geopotential height.
A	wave amplitude.
B_ν	Planck function for radiation of frequency ν.
D	Diffusion tensor for zonal=mean tracer transport.
F	Irradiance (Flux density).
F_ν	Monochromatic Irradiance.
H	Pressure scale height (7km).
I	Radiance.
I_ν	Monochromatic Radiance.
J	radiative heating rate ($Jkg^{-1}s^{-1}$).
J_ν	Monochromatic source function.
K	Zonal-mean tracer transport tensor.
L	Advective (antisymmetric) part of K.
M	Montgomery Potential.

N	buoyancy frequency.
P	$\log(p_o/p)$
Q	potential vorticity.
R	Gas constant for dry air ($287 Jkg^{-1}K^{-1}$).
S	photochemical source term.
T	temperature.
T_e	Radiative equilibrium temperature.
T_o	Standard zonal-mean temperature profile.
$\mathbf{U}=(U,V)=(\sigma_\eta u,\sigma_\eta v)$	horizontal flux of matter.
$\mathbf{U}_D$	Divergent part of **U**.
$W=\sigma_\eta w$	vertical flux of matter.
$\mathbf{X}=(X,Y)$	External force per unit mass.
α	Inverse dissipation time for infrared damping.
β	Inverse dissipation time for photochemical acceleration.
ϕ	latitude; azimuthal angle.
ε	error.
λ	longitude.
$\eta=\log\theta$	isentropic vertical coordinate.
κ	$R/c_p(\simeq 2/7)$.
μ	tracer mixing ratio; cosine of zenith angle.
ν	frequency of electromagnetic radiation.
ρ	density in height coordinates.
ρ_a	constituent density.
θ	potential temperature; zenith angle.
σ	Stefan-Boltzman constant.
σ_η	density in isentropic coordinates.
τ	timescale; optical depth.

ξ	perturbation displacement.
χ	potential function for divergent horizontal mass flux.
ζ	vertical component of vorticity.
Ψ	streamfunction for rotational mass flux.
$\Lambda=(\Lambda_\phi,\Lambda_\eta)$	eddy flux of tracer in the meridional plane.
Ω	angular frequency of Earth's rotation.
Γ	Static Stability.
Γ_e	Static Stability at radiative equilibrium.
Υ	transmittance.
$\Pi=(\Pi_\lambda,\Pi_\phi)$	generalised potential vorticity flux vector.

I.III. Operators.

∇_η	Horizontal del operator at constant entropy.
∇_M	Del operator in meridional plane.
D_η	Material derivative at constant entropy $(\partial/\partial t + \mathbf{u}\cdot\nabla_\eta)$.
D_Z	Material derivative following the zonal-mean flow $(\partial/\partial t + (a\cos\phi)^{-1}[u]\partial/\partial\lambda)$.

References.

Al-Ajmi, D.N., Harwood, R.S. and Miles, T., 1985; "A Sudden Warming in the Middle Atmosphere of the Southern Hemisphere."; Quart. J. Royal Met. Soc., **111**, pp359-390.

Andrews, D.G., 1983; "A Finite Amplitude Eliassen-Palm Flux Theorem in Isentropic Coordinates."; J. Atmos. Sci., **41**, pp1877-1883.

Andrews, D.G., 1985; "Wave - Mean-Flow Interaction in the Stratosphere."; Adv. Geophys., **28A**, pp149-275.

Andrews, D.G., 1987; "On the Interpretation of the Eliassen-Palm Flux Divergence."; Quart. J. Royal Met. Soc., **113**, pp323-338.

Andrews, D.G. and McIntyre, M.E., 1976; "Planetary Waves in Horizontal and Vertical Shear: the Generalised Eliassen-Palm Relation and the Mean Zonal Acceleration."; J. Atmos. Sci., **33**, pp2031-2048.

Andrews, D.G. and McIntyre, M.E., 1978; "Generalised Eliassen-Palm and Charney-Drazin Theorems for Waves on Axisymmetric Mean Flows in Compressible Atmospheres."; J. Atmos. Sci., **35**, pp175-185.

Barnett, J.J. and Corney, M., 1985; "Middle Atmosphere Reference Model Derived From Satellite Data."; pp47-85 in Handbook for MAP, **16**: *Atmospheric Structure and its Variation in the Region 20-120km: Draft of a New Reference Middle Atmosphere.* Edited by K. Labitzke, J.J. Barnett and B. Edwards.

Barnett, J.J., Houghton, J.T. and Pyle, J.A., 1975; "The Temperature Dependence of the Ozone Concentration near the Stratopause."; Quart. J. Royal Met. Soc., **101**, pp245-257.

Blake, D. and Lindzen, R.S., 1973; "The Effect of Photochemical Models on Calculated equilibria and Cooling Rates in the Stratosphere."; Mon. Weath. Rev., **101**, pp783-802.

Boville, B.A., 1987; "The validity of the geostrophic approximation in the winter stratosphere and troposphere."; J. Atmos. Sci., **44**, pp443-457.

Brasseur, G. and Solomon, S., 1986; "Aeronomy of the Middle Atmosphere."; D. Reidel Publ. Co., (2nd Ed.), 452pp.

Butchart, N. and Remsberg, E.E., 1986; "The Area of the Stratospheric Polar Vortex as a Diagnostic for Tracer Transport on Isentropic Surfaces."; J. Atmos. Sci., **43**, pp1319-1339.

Callis, L.B., Boughner, R.E. and Lambeth, J.D., 1987; "The Stratosphere: Climatologies of Radiative Heating and Cooling Rates and the Diabatically Diagnosed Net Circulation Fields."; J. Geophys. Res., **92D**, pp5585-5607.

Chapman, S., 1931; "A Theory of Upper Atmospheric Ozone."; Royal Met. Soc. Mem., **3**, pp103-125.

Charney, J.G. and Drazin, P.G., 1961; "Propagation of Planetary Scale Disturbances from the Lower into the Upper Stratosphere."; J. Geophys. Res., **66**, pp83-109.

Charney, J.G. and Stern, M.E., 1962; "On the Stability of Baroclinic Internal Baroclinic Jets in a Rotating Atmosphere."; J. Atmos. Sci., **19**, pp159-172.

Clough, S.A., Grahame, N.S. and O'Neill, A., 1985; "Potential Vorticity in the Stratosphere Derived Using Data From Satellites."; Quart. J. Royal Met. Soc., **111**, pp335-358.

Coy, L. and Hitchman, M., 1984; "Kelvin Wave Packets and Flow Acceleration: a Comparison of Modelling and Observation."; J. Atmos. Sci., **41**, pp1875-1880.

Danielsen, E.F., 1959; "The Laminar Structure of the Atmosphere and its Relationship to the Concept of a Tropopause"; Arch. Met. Geosphys. Biokl., **A11**, pp292-332.

Dickinson, R.E., 1969; "Theory of Planetary Wave-Zonal Flow Interaction."; J. Atmos. Sci., **26**, pp73-81.

Dickinson, R.E., 1973; "Methods of Parameterisation for Infrared Cooling Between Altitudes of 30 and 70km."; J. Geophys. Res., **78**, pp4451-4457.

Dunkerton, T.J., 1978; "On the Mean Meridional Mass Motions of the Stratosphere and Mesosphere."; J. Atmos. Sci., **35**, pp2325-2333.

Dunkerton, T.J. and Delisi, D.P., 1986; "Evolution of Potential Vorticity in the Winter Stratosphere of January-February 1979."; J. Geophys. Res., **91D**, pp1199-1208.

Dutton, J.A., 1976; "*The Ceaseless Wind:* an Introduction to the Theory of Atmospheric Motion."; McGraw-Hill Inc, pp579.

Eliassen, A.N. and Palm, E., 1961; "On the Transfer of Energy in Stationary Mountain Waves."; Geofys. Publikasjoner, **22**, pp1-22.

Ertel, H., 1942; "Ein Nueur Hydrodynamischer Wirbelsatz."; Meteor. Z., **59**, pp277-281.

Farman, J.C., Gardiner, B.G. and Shanklin, J.D., 1985; "Large Losses of Total Ozone in Antarctic Reveal Seasonal ClO_x/NO_x Interaction."; Nature, **35**, pp207-210.

Fels, S.B., 1982; "A Parameterisation of Scale-Dependent Radiative Damping Rates in the Middle Atmosphere."; J. Atmos. Sci., **39**, 1141-1152

Fels, S.B., 1985; "Radiative-Dynamical Interactions in the Middle Atmosphere."; Adv. Geophys., **28A**, pp277-300.

Gallimore, R.G. and Johnson, D.R., 1981; "The Forcing of the Meridional Circulation of the Isentropic Zonally Averaged Circumpolar Vortex"; J. Atmos. Sci., **38**, pp583-599.

Ghazi, A., Ramanathan, V. and Dickinson, R.E., 1979; "Acceleration of Upper Stratospheric Radiative Damping: Observational Evidence."; Geophys. Res. Lett., **6**, pp437-440.

Ghazi, A., Wang, P-H. and McCormick, M.P., 1985; "A Study on Radiative Damping of Planetary Waves."; J. Atmos. Sci., **42**, pp2032-2042.

Gille, J.C. and Lyjak, L.V., 1986; "Radiative Heating and Cooling Rates in the Middle Atmosphere."; J. Atmos. Sci., **43**, 2215-2229

Gille, J.C., Lyjak, L.V. and Smith, A.K., 1987; "The Global Residual Mean Circulation in the Middle Atmosphere for the Northern Winter Period."; J. Atmos. Sci., **44**, pp1437-1452.

Gille, J.C. and Russell, J.M. III, 1984; "The Limb Infrared Monitor of the Stratosphere: Experiment Description, performance and Results."; J. Geophys. Res., **89**, 5125-5140

Gille, J.C., Russell, J.M. III, Bailey, P.L., Gordley, L.L., Remsberg, E.E., Liensesch, J.H., Planet, W.G., House, F.G., Lyjak, L.V. and Beck, S.A., 1984a; "Validation of Temperature Retrievals Obtained by the Limb Infrared Monitor of the Stratosphere (LIMS) Experiment on NIMBUS 7."; J. Geophys. Res., **89D4**, pp5147-5160.

Gille, J.C., Russell, J.M. III, Bailey, P.L., Remsberg, E.E., Gordley, L.L., Evans, W.F.J., Fischer, H., Gandrud, B.W., Girard, A., Harries, J.E. and Beck, S.A., 1984b; "Accuracy and Precision of the Nitric Acid Concentrations Determined by the LIMS Experiment on Nimbus 7."; J. Geophys. Res., **89D**, pp5179-5190.

Gray, L.J. and Pyle, J.A., 1987; "Two Dimensional Model Studies of Equatorial Dynamics and Tracer Distributions."; Quart. J. Royal Met. Soc., **113**, pp637-654.

Haigh, J.D., 1984; "Radiative Heating in the Lower Stratosphere and the Distribution of Ozone in a Two-Dimensional Model"; Quart. J. Royal Met. Soc., **110**, pp167-185.

Haigh, J.D., 1985; "A Fast Method for Calculating Scale-Dependent Photochemical Acceleration in Dynamical Models of the Stratosphere."; Quart. J. Royal Met. Soc., **111**, 1027-1038

Hartmann, D.L., 1978; "A Note Concerning the Effect of Varying Extinction on Radiative-Photochemical Acceleration"; J. Atmos. Sci., **35**, pp1125-1130.

Hartmann, D.L., 1981; "Some Aspects of the Coupling Between Radiation, Chemistry and Dynamics in the Stratosphere"; J. Geophys. Res., **86**, pp9631-9640.

Harwood, R.S., 1980; "Dynamical Models of the Middle Atmosphere for Tracer Studies."; Phil. Trans. R. Soc. Lond., **A296**, pp103-127.

Haynes, P.H. and McIntyre, M.E., 1987; "On the Evolution of Vorticity and Potential Vorticity in the Presence of Nonconservative Forces."; J. Atmos. Sci., **44**, pp828-841.

Henderson-Sellers, A., 1986; "Layer Cloud Amounts For January and July 1979 From 3D-Nephanalysis."; J. Clim. Appl. Met., **25**, pp118-132.

Hitchman, M.H. and Leovy, C.B., 1986; "Evolution of the Zonal Mean State in the Equatorial Middle Atmosphere During October 1978 - May 1979."; J. Atmos. Sci., **43**, pp3159-3176.

Holton, J.R., 1986; "Meridional Distribution of Stratospheric Trace Constituents."; J. Atmos. Sci., **43**, pp1238-1242.

Holton, J.R. and Lindzen, R.S., 1972; "An Updated Theory of the Quasi-Biennial Oscillation of the Tropical Stratosphere."; J. Atmos. Sci., **29**, pp1076-1080.

Houghton, D.D., 1968; "Derivation of the Elliptic Condition for the Balance Equation in Spherical Coordinates."; J. Atmos. Sci., **25**, pp927-928.

Houghton, J.T., 1978; "The Physics of Atmospheres."; Cambridge University Press, 203pp.

Jones, R.L. and Pyle, J.A., 1984; "Observations of CH_4 and N_2O by the NIMBUS-7 SAMS: A Comparison With In Situ Data and Two Dimensional Numerical Model Calculations."; J. Geophys. Res., **D89**, pp5263-5279.

Juckes, M.N. and McIntyre, M.E., 1987; "A High-Resolution One-Layer Model of Breaking Planetary Waves in the Stratosphere"; Nature, **328**, pp590-596.

Keating, G.M. and Young, D.F., 1985; "Interim Reference Ozone Models for the Middle Atmosphere."; pp205-209 in Handbook for MAP, **16**: *Atmospheric Structure and its Variation in the Region 20-120km: Draft of a New Reference Middle Atmosphere.* Edited by K. Labitzke, J.J. Barnett and B. Edwards.

Kiehl, J.T. and Solomon, S., 1986; "On the Radiative Balance of the Stratosphere."; J. Atmos. Sci., **43**, pp1525-1534.

Killworth, P.D. and McIntyre, M.E., 1985; "Do Rossby Wave Critical Layers Absorb, Reflect of Over-reflect?"; J. Fluid Mech., **161**, pp449-492.

Ko, M.K., Tung, K.-K., Weisenstein, D.K. and Sze, N.D., 1985; "A Zonal Mean Model of Stratospheric Tracer Transport in Isentropic Coordinates: Numerical Simulations for Nitrous Oxide and Nitric Acid."; J. Geophys. Res., **90D**, pp2313-2329.

Labitzke, K., 1981; "The amplification of Height Wave 1 in January 1979: A Characteristic Precondition for the Major Warming in February."; Mon. Weath. Rev., **109**, pp983-989.

Lacis, A.A. and Hansen, J.E., 1974; "A Parameterisation for the Absorption of Solar Radiation in the Earth's Atmosphere."; J. Atmos. Sci., **31**, pp118-133.

Leovy, C.B., Sun, C.-R., Hitchman, M.H., Remsberg, E.E., Russell, J.M. III, Gordley, L.L., Gille, J.C. and Lyjak, L.V., 1985; "Transport of Ozone in the Middle Stratosphere: Evidence for Planetary Wave Breaking."; J. Atmos. Sci., **42**, pp230-244.

Lindzen, R.S. and Goody, R., 1965; "Radiative and Photochemical Processes in Mesospheric Dynamics: Part 1, Models for Radiative and Photochemical Processes."; J. Atmos. Sci., **22**, pp341-348.

Lindzen, R.S. and Will, D.I., 1973; "An Analytic Formula for Heating Rate Due to Ozone Absorption."; J. Atmos. Sci., **30**, pp513-515.

London, J., 1957; "A Study of the Atmospheric Heat Balance."; New York University, Report AF19, pp122-165.

Lorenz, E.N., 1967; "The Nature and Theory of the General Circulation of the Atmosphere."; World Meteorological Organisation, **No 218, TP.115**, 158pp.

Madden, R.A. and Labitzke, K., 1981; "A Free Rossby Wave in the Troposphere and Stratosphere During January 1979."; J. Geophys. Res., **86C**, pp1247-1254.

Manabe, S. and Wetherald, R.T., 1967; "Thermal Equilibrium of the Atmosphere With a Given Distribution of Relative Humidity."; J. Atmos. Sci., **24**, pp241-259.

Matsuno, T., 1970; "Vertical Propagation of Stationary Planetary Waves in the Winter Northern Hemisphere."; J. Atmos. Sci., **27**, pp871-883.

McIntyre, M.E. and Palmer, T.N., 1983; "Breaking Planetary Waves in the Stratosphere."; Nature, **305**, pp593-600.

McIntyre, M.E. and Palmer, T.N., 1984; "The 'Surf Zone' in the Stratosphere."; J. Atmos. Terr. Phys, **46**, pp825-849.

McPeters, R.D., Heath, P.F. and Bhartia, P.K., 1984; "Average Ozone Profiles from the Nimbus 7 SBUV Instrument."; J. Geophys. Res., **89**, pp5199-5214.

McWilliams, J.C. and Gent, P.R., 1980; "Intermediate Models of Planetary Circulations in the Atmosphere and Oceans."; J. Atmos. Sci., **37**, pp1657-1678.

Murgatroyd, R.J. and Singleton, F., 1961; "Possible Meridional Circulations in the Stratosphere and Mesosphere."; Quart. J. Royal Met. Soc., **87**, pp125-135.

Plumb, R.A. and Mahlman, J.D., 1987; "The Zonally-Averaged Transport Characteristics of the GFDL General Circulation Transport Model."; J. Atmos. Sci., **44**, pp298-327.

Ramanathan, V., 1976; "Radiative Transfer Within the Earth's Troposphere and Stratosphere: a Simplified Radiative-Convective Model."; J. Atmos. Sci., **33**, pp1330-1346.

Reed, R.J. and German, D., 1965; "A Contribution to the Problem of Stratospheric Diffusion by Large-Scale Mixing."; Mon. Weath. Rev., **93**, pp313-321.

Remsberg, E.E., Russell, J.M. III, Gille, J.C., Gordley, L.L., Bailey, P.L., Planet, W.G. and Harries, J.E., 1984a; "The Validation of NIMBUS 7 LIMS

Measurements of Ozone."; J. Geophys. Res., **89D4**, pp5161-5178.

Remsberg, E.E., Russell, J.M.III, Gordley, L.L., Gille, J.C. and Bailey, P.L., 1984b; "Implications of the Stratospheric Water Vapour Distribution as Determined from the Nimbus 7 LIMS Experiment."; J. Atmos. Sci., **41**, pp2934-2945.

Robinson, W.A., 1986; "The Application of the Quasi-geostrophic Eliassen-Palm Flux to the Analysis of Stratospheric Data."; J. Atmos. Sci., **43**, pp1017-1023.

Rodgers, C.D., 1967; "The Radiative Heat Budget of the Troposphere and Lower Stratosphere."; MIT Planetary Circulation Project, Report A2.

Rodgers, C.D. (*ed.*), 1984; "Coordinated Study of the Behaviour of the Middle Atmosphere in Winter."; Handbook for MAP, Volume 12, 154pp.

Rodgers, C.D. and Walshaw, C.D., 1966; "The Computation of Infrared Cooling Rate in Planetary Atmospheres."; Quart. J. Royal Met. Soc., **92**, pp67-92.

Rosenfield, J.E., Schoerberl, M.R. and Geller, M.A., 1987; "A Computation of the Stratospheric Diabatic Circulation Using an Accurate Radiative Transfer Model."; J. Atmos. Sci., **44**, pp859-876.

Russell, J.M. III (*ed.*), 1986; "Middle Atmosphere Composition Revealed by Satellite Observations."; Handbook for MAP, Volume 22. 302pp.

Russell, J.M. III, Gille, J.C., Remsberg, E.E., Gordley, L.L., Bailey, P.L., Drayson, S.R., Fischer, H., Girard, A., Harries, J.E. and Evans, W.F.J., 1984a; "Validation of Nitrogen Dioxide Results Measured by the LIMS Experiment on Nimbus 7."; J. Geophys. Res., **89D**, pp5099-5107.

Russell, J.M. III, Gille, J.C., Remsberg, E.E., Gordley, L.L., Bailey, P.L., Fischer, H., Girard, A., Drayson, S.R., Evans, W.F.J. and Harries, J.E., 1984b; "Validation of Water Vapour Results Measured by the LIMS Experiment on NIMBUS 7."; J. Geophys. Res., **89D4**, pp5115-5124.

Salby, M.L., Hartmann, D.L., Bailey, P.L. and Gille, J.C., 1984; "Evidence for Equatorial Kelvin Modes in Nimbus 7 LIMS."; J. Atmos. Sci., **41**, pp220-235.

Sawyer, J.S., 1964; "Dynamical Aspects of the Bi-Polar Stratospheric Circulation."; Quart. J. Royal Met. Soc., **90**, pp395-404.

Shine, K.P., 1987; "The Middle Atmosphere in the Absence of Dynamical Heat Fluxes."; Quart. J. Royal Met. Soc., **113**, pp605-636.

Solomon, S., Kiehl, J.T., Garcia, R.R. and Grose, W., 1987; "Tracer Transport by the Diabatic Circulation Deduced From Satellite Observations."; J. Atmos. Sci., **43**, pp1603-1617.

Smith, A.K. and Bailey, P.L., 1985; "Comparison of Horizontal Winds From the LIMS Satellite Instrument With Rocket Measurements."; J. Geophys. Res., **90D**, pp3897-3901.

Starr, V.P., 1945; "A Quasi-Lagrangian System of Hydrodynamical Equations."; J. Meteor., **2**, pp227-237.

Strobel, D.F., 1977; "Photochemical-Radiative Damping and Instability in the Stratosphere."; Geophys. Res. Lett, **4**, pp424-426.

Strobel, D.F., 1978; "Parameterisation of the Atmospheric Heating Rate From 15-120km due to O_2 and O_3 Absorption of Solar Radiation."; J. Geophys. Res., **83**, 6225-6230

Tung, K-K., 1982; "On the Two-Dimensional Transport of Stratospheric Trace Gases in Isentropic Coordinates."; J. Atmos. Sci., **39**, pp2330-2355.

Tung, K-K., 1986; "Nongeostrophic Theory of Zonally Averaged Circulation. Part 1: Formulation."; J. Atmos. Sci., **43**, pp2600-2618.

World Meteorological Association, 1986; "Atmospheric Ozone 1985, Assessment of our understanding of the processes controlling its present distribution and changes."; WMO Report No 16.